

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
29 November 2001 (29.11.2001)

PCT

(10) International Publication Number  
**WO 01/90358 A2**

(51) International Patent Classification<sup>7</sup>: C12N 15/12,  
C07K 14/715, 16/18, G01N 33/53, C12N 5/10

(21) International Application Number: PCT/US01/16767

(22) International Filing Date: 23 May 2001 (23.05.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/206,862 24 May 2000 (24.05.2000) US

(71) Applicant: SCHERING CORPORATION [US/US];  
2000 Galloping Hill Road, Kenilworth, NJ 07033-0530  
(US).

(72) Inventor: GORMAN, Daniel, M.; 6371 Central Avenue,  
Newark, CA 94560 (US).

(74) Agent: ZARADIC, Sandy; Schering-Plough Corpora-  
tion, Patent Department, K-6-1, 1990, 2000 Galloping Hill  
Road, Kenilworth, NJ 07033-0530 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CZ,  
DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, HR, HU,  
ID, IL, IN, IS, JP, KG, KR, KZ, LC, LK, LR, LT, LU, LV,  
MA, MD, MG, MK, MN, MX, MZ, NO, NZ, PL, PT, RO,  
RU, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UZ, VN,  
YU, ZA.

(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Declaration under Rule 4.17:**

— *as to the applicant's entitlement to claim the priority of the  
earlier application (Rule 4.17(iii)) for all designations*

**Published:**

— *without international search report and to be republished  
upon receipt of that report*

*For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*



WO 01/90358 A2

(54) Title: MAMMALIAN RECEPTOR PROTEINS; RELATED REAGENTS AND METHODS

(57) Abstract: Nucleic acids encoding mammalian, e.g., primate, receptors, purified receptor proteins and fragments thereof. Anti-  
bodies, both polyclonal and monoclonal, are also provided. Methods of using the compositions for both diagnostic and therapeutic  
utilities are described.

5

## MAMMALIAN RECEPTOR PROTEINS; RELATED REAGENTS AND METHODS

10

## FIELD OF THE INVENTION

The present invention relates to compositions and methods for affecting mammalian physiology, including immune system function. In particular, it provides methods to regulate development and/or the immune system. Diagnostic and therapeutic uses of these materials are also disclosed.

15

## BACKGROUND OF THE INVENTION

Recombinant DNA technology refers generally to techniques of integrating genetic information from a donor source into vectors for subsequent processing, such as through introduction into a host, whereby the transferred genetic information is copied and/or expressed in the new environment. Commonly, the genetic information exists in the form of complementary DNA (cDNA) derived from messenger RNA (mRNA) coding for a desired protein product. The carrier is frequently a plasmid having the capacity to incorporate cDNA for later replication in a host and, in some cases, actually to control expression of the cDNA and thereby direct synthesis of the encoded product in the host. See, e.g., Sambrook, et al. (1989) Molecular Cloning: A Laboratory Manual, (2d ed.) vols. 1-3, CSH Press, NY.

25

For some time, it has been known that the mammalian immune response is based on a series of complex cellular interactions, called the "immune network". Recent research has provided new insights into the inner workings of this network. While it remains clear that much of the immune response does, in fact, revolve around the network-like interactions of lymphocytes, macrophages, granulocytes, and other cells, immunologists now generally hold the opinion that soluble proteins, known as lymphokines, cytokines, or monokines, play critical roles in controlling these cellular interactions. Thus, there is considerable interest in the isolation, characterization, and mechanisms of action of cell modulatory factors, an understanding of which will lead to significant advancements in the diagnosis and therapy of numerous medical abnormalities, e.g., immune system disorders.

30

35



The immune system of vertebrates consists of a number of organs and several different cell types. Two major cell types include the myeloid and lymphoid lineages. Among the lymphoid cell lineage are B cells, which were originally characterized as differentiating in fetal liver or adult bone marrow, and T cells, which were originally  
5 characterized as differentiating in the thymus. See, e.g., Paul (ed. 1998) Fundamental Immunology (4th ed.) Raven Press, New York; and Thomson (ed. 1994) The Cytokine Handbook 2d ed., Academic Press, San Diego. Lymphokines apparently mediate cellular activities in a variety of ways. They have been shown to support the proliferation, growth, and/or differentiation of cells, e.g., pluripotential hematopoietic stem cells, into  
10 vast numbers of progenitors comprising diverse cellular lineages which make up a complex immune system. Proper and balanced interactions between the cellular components are necessary for a healthy immune response. The different cellular lineages often respond in a different manner when lymphokines are administered in conjunction with other agents.

15 Cell lineages especially important to the immune response include two classes of lymphocytes: B-cells, which can produce and secrete immunoglobulins (proteins with the capability of recognizing and binding to foreign matter to effect its removal), and T-cells of various subsets that secrete lymphokines and induce or suppress the B-cells and various other cells (including other T-cells) making up the immune network. These  
20 lymphocytes interact with many other cell types.

Research to better understand and treat various immune disorders has been hampered by the general inability to maintain cells of the immune system in vitro. Immunologists have discovered that culturing many of these cells can be accomplished through the use of T-cell and other cell supernatants, which contain various growth  
25 factors, including many of the lymphokines.

Various growth and regulatory factors exist which modulate morphogenetic development. And many receptors for cytokines are also known. Often there are at least two critical subunits in the functional receptor. See, e.g., Gonda and D'Andrea (1997) Blood 89:355-369; Presky, et al. (1996) Proc. Nat'l Acad. Sci. USA 93:14002-14007;  
30 Drachman and Kaushansky (1995) Curr. Opin. Hematol. 2:22-28; Theze (1994) Eur. Cytokine Netw. 5:353-368; and Lemmon and Schlessinger (1994) Trends Biochem. Sci. 19:459-463.

From the foregoing, it is evident that the discovery and development of new soluble proteins and their receptors, including ones similar to lymphokines, should  
35 contribute to new therapies for a wide range of degenerative or abnormal conditions which directly or indirectly involve development, differentiation, or function, e.g., of the

immune system and/or hematopoietic cells. In particular, the discovery and understanding of novel receptors for lymphokine-like molecules which enhance or potentiate the beneficial activities of other lymphokines would be highly advantageous. However, the lack of understanding of how the immune system is regulated or differentiates has blocked the ability to advantageously modulate the normal defensive mechanisms to biological challenges. Medical conditions characterized by abnormal or inappropriate regulation of the development or physiology of relevant cells thus remain unmanageable. The discovery and characterization of specific cytokines and their receptors will contribute to the development of therapies for a broad range of degenerative or other conditions which affect the immune system, hematopoietic cells, as well as other cell types. The present invention provides new receptors for ligands exhibiting similarity to cytokine like compositions and related compounds, and methods for their use.

#### SUMMARY OF THE INVENTION

The present invention is directed to novel receptors related to cytokine receptors, e.g., primate, cytokine receptor like molecular structures, designated DNAX Cytokine Receptor Subunits (DCRS), and their biological activities. In particular, it provides description of various subunits, designated DCRS6, DCRS7, DCRS8, DCRS9, and DCRS10. Primate, e.g, human, and rodent, e.g., mouse, embodiments of the various subunits are provided. It includes nucleic acids coding for the polypeptides themselves and methods for their production and use. The nucleic acids of the invention are characterized, in part, by their homology to cloned complementary DNA (cDNA) sequences enclosed herein.

The present invention provides a composition of matter selected from: a substantially pure or recombinant polypeptide comprising at least three distinct nonoverlapping segments of at least four amino acids identical to segments of SEQ ID NO: 2, 5, 8, 11, 23, or 26; a substantially pure or recombinant polypeptide comprising at least three distinct nonoverlapping segments of at least four amino acids identical to segments of SEQ ID NO: 14; a substantially pure or recombinant polypeptide comprising at least two distinct nonoverlapping segments of at least five amino acids identical to segments of SEQ ID NO: 14; a natural sequence DCRS8 comprising mature SEQ ID NO: 14; a fusion polypeptide comprising DCRS8 sequence; a substantially pure or recombinant polypeptide comprising at least three distinct nonoverlapping segments of at least four amino acids identical to segments of SEQ ID NO: 17 or 20; a substantially pure or recombinant polypeptide comprising at least two distinct nonoverlapping segments of at least five amino acids identical to segments of SEQ ID NO: 17 or 20; a natural

sequence DCRS9 comprising mature SEQ ID NO: 17 or 20; or a fusion polypeptide comprising DCRS9 sequence. Preferably, wherein the distinct nonoverlapping segments of identity include: one of at least eight amino acids; one of at least four amino acids and a second of at least five amino acids; at least three segments of at least four, five, and six amino acids, or one of at least twelve amino acids. In other embodiments, the:  
5 polypeptide: comprises a mature sequence of Tables 1, 2, 3, 4, or 5; is an unglycosylated form of DCRS8 or DCRS9; is from a primate, such as a human; comprises at least seventeen amino acids of SEQ ID NO: 14 or 17; exhibits at least four nonoverlapping segments of at least seven amino acids of SEQ ID NO: 14 or 17; is a natural allelic  
10 variant of DCRS8 or DCRS9; has a length at least about 30 amino acids; exhibits at least two non-overlapping epitopes which are specific for a primate DCRS8 or DCRS9; is glycosylated; has a molecular weight of at least 30 kD with natural glycosylation; is a synthetic polypeptide; is attached to a solid substrate; is conjugated to another chemical moiety; is a 5-fold or less substitution from natural sequence; or is a deletion or insertion  
15 variant from a natural sequence.

The invention further embraces a composition comprising: a substantially pure DCRS8 or DCRS9 and another cytokine receptor family member; a sterile DCRS8 or DCRS9 polypeptide; the DCRS8 or DCRS9 polypeptide and a carrier, wherein the carrier is: an aqueous compound, including water, saline, and/or buffer; and/or formulated for  
20 oral, rectal, nasal, topical, or parenteral administration. Additional embodiments include a polypeptide comprising: mature protein sequence of Tables 1, 2, 3, 4, or 5; a detection or purification tag, including a FLAG, His6, or Ig sequence; or sequence of another cytokine receptor protein. Kit embodiments include ones comprising a described polypeptide, and: a compartment comprising the protein or polypeptide; or instructions  
25 for use or disposal of reagents in the kit.

Binding compositions are provided, e.g., comprising an antigen binding site from an antibody, which specifically binds to a natural DCRS8 or DCRS9 polypeptide, wherein: the binding compound is in a container; the DCRS8 or DCRS9 polypeptide is from a human; the binding compound is an Fv, Fab, or Fab2 fragment; the binding  
30 compound is conjugated to another chemical moiety; or the antibody: is raised against a peptide sequence of a mature polypeptide of Table 3 or 4; is raised against a mature DCRS8 or DCRS9; is raised to a purified human DCRS8 or DCRS9; is immunoselected; is a polyclonal antibody; binds to a denatured DCRS8 or DCRS9; exhibits a K<sub>d</sub> to antigen of at least 30 μM; is attached to a solid substrate, including a bead or plastic membrane; is  
35 in a sterile composition; or is detectably labeled, including a radioactive or fluorescent label. Kits include ones comprising such a binding compound, and: a compartment

comprising the binding compound; or instructions for use or disposal of reagents in the kit.

The invention also provides methods of producing an antigen:antibody complex, comprising contacting under appropriate conditions a primate DCRS8 or DCRS9 polypeptide with a described antibody, thereby allowing the complex to form. Preferred methods include ones wherein: the complex is purified from other cytokine receptors; the complex is purified from other antibody; the contacting is with a sample comprising an interferon; the contacting allows quantitative detection of the antigen; the contacting is with a sample comprising the antibody; or the contacting allows quantitative detection of the antibody. Further compositions include those comprising: a sterile binding compound, as described, or the binding compound and a carrier, wherein the carrier is: an aqueous compound, including water, saline, and/or buffer; and/or formulated for oral, rectal, nasal, topical, or parenteral administration.

Nucleic acid compositions include an isolated or recombinant nucleic acid encoding a described polypeptide wherein the: DCRS8 or DCRS9 is from a human; or the nucleic acid: encodes an antigenic peptide sequence of Table 3 or 4; encodes a plurality of antigenic peptide sequences of Table 3 or 4; exhibits identity over at least thirteen nucleotides to a natural cDNA encoding the segment; is an expression vector; further comprises an origin of replication; is from a natural source; comprises a detectable label; comprises synthetic nucleotide sequence; is less than 6 kb, preferably less than 3 kb; is from a primate; comprises a natural full length coding sequence; is a hybridization probe for a gene encoding the DCRS8 or DCRS9; or is a PCR primer, PCR product, or mutagenesis primer. Also provided are a cell or tissue comprising such a recombinant nucleic acid, e.g., where the cell is: a prokaryotic cell; a eukaryotic cell; a bacterial cell; a yeast cell; an insect cell; a mammalian cell; a mouse cell; a primate cell; or a human cell.

Kit embodiments include those comprising a described nucleic acid and: a compartment comprising the nucleic acid; a compartment further comprising a primate DCRS8 or DCRS9 polypeptide; or instructions for use or disposal of reagents in the kit.

Other nucleic acids provided include ones which: hybridize under wash conditions of 30 minutes at 30° C and less than 2M salt to the coding portion of SEQ ID NO: 13 or 16; or exhibit identity over a stretch of at least about 30 nucleotides to a primate DCRS8 or DCRS9. Preferably, such will be nucleic acids where: the wash conditions are: at 45° C and/or 500 mM salt; at 55° C and/or 150 mM salt; or the stretch is at least 55 or 75 nucleotides.

Also provided are methods of modulating physiology or development of a cell or tissue culture cells comprising contacting the cell with an agonist or antagonist of a

mammalian DCRS8 or DCRS9. Preferably, the cell is transformed with a nucleic acid encoding the DCRS8 or DCRS9 and another cytokine receptor subunit.

## DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

5

### OUTLINE

- I. General
- II. Activities
- III. Nucleic acids
  - 10 A. encoding fragments, sequence, probes
  - B. mutations, chimeras, fusions
  - C. making nucleic acids
  - D. vectors, cells comprising
- 15 IV. Proteins, Peptides
  - A. fragments, sequence, immunogens, antigens
  - B. muteins
  - C. agonists/antagonists, functional equivalents
  - D. making proteins
- 20 V. Making nucleic acids, proteins
  - A. synthetic
  - B. recombinant
  - C. natural sources
- 25 VI. Antibodies
  - A. polyclonals
  - B. monoclonal
  - C. fragments; Kd
  - D. anti-idiotypic antibodies
  - E. hybridoma cell lines
- 30 VII. Kits and Methods to quantify DCRSs
  - A. ELISA
  - B. assay mRNA encoding
  - C. qualitative/quantitative
  - D. kits
- 35 VIII. Therapeutic compositions, methods
  - A. combination compositions
  - B. unit dose
  - C. administration
- IX. Screening
- X. Ligands

40

### I. General

45 The present invention provides the amino acid sequence and DNA sequence of mammalian, herein primate, cytokine receptor-like subunit molecules, these designated DNAX Cytokine Receptor Subunits 6 (DCRS6), 7 (DCRS7), 8 (DCRS8), 9 (DCRS9), and 10 (DCRS10) having particular defined properties, both structural and biological.

Various cDNAs encoding these molecules were obtained from primate, e.g., human, and/or rodent, e.g., mouse, cDNA sequence libraries. Other primate or other mammalian counterparts would also be desired.

Some of the standard methods applicable are described or referenced, e.g., in  
 5 Maniatis, et al. (1982) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor Press; Sambrook, et al. (1989) Molecular Cloning: A Laboratory Manual, (2d ed.), vols. 1-3, CSH Press, NY; Ausubel, et al., Biology, Greene Publishing Associates, Brooklyn, NY; or Ausubel, et al. (1987 and periodic supplements) Current Protocols in Molecular Biology, Greene/Wiley, New York; each of which is  
 10 incorporated herein by reference.

Nucleotide (SEQ ID NO: 1) and corresponding amino acid sequence (SEQ ID NO: 2) of a primate, e.g., human, DCRS6 coding segment is shown in Table 1 along with reverse translation (SEQ ID NO: 3). Rodent, e.g., mouse, counterpart sequences are provided, e.g., SEQ ID NO: 4-6.

15 Similarly, nucleotide (SEQ ID NO: 7) and corresponding amino acid sequence (SEQ ID NO: 8) of a primate, e.g., human, DCRS7 coding segment is shown in Table 2 along with reverse translation (SEQ ID NO: 9). Rodent, e.g., mouse, counterpart sequences are provided, e.g., SEQ ID NO: 10-12. Nucleotide (SEQ ID NO: 13) and corresponding amino acid sequence (SEQ ID NO: 14) of a primate, e.g., human, DCRS8  
 20 coding segment is shown in Table 3 along with reverse translation (SEQ ID NO: 15).

Nucleotide (SEQ ID NO: 16) and corresponding amino acid sequence (SEQ ID NO: 17) of a primate, e.g., human, DCRS9 coding segment is shown in Table 4 along with reverse translation (SEQ ID NO: 18). Rodent, e.g., mouse, counterpart sequences are provided, e.g., SEQ ID NO: 19-21. Nucleotide (SEQ ID NO: 22) and corresponding  
 25 amino acid sequence (SEQ ID NO: 23) of a primate, e.g., human, DCRS10 coding segment is shown in Table 5 along with reverse translation (SEQ ID NO: 24). Rodent, e.g., mouse, counterpart sequences are provided, e.g., SEQ ID NO: 26-27.

30 Table 1: Nucleotide and polypeptide sequences of DNAX Cytokine Receptor Subunit like embodiments (DCRS6). Primate, e.g., human, embodiment (see SEQ ID NO: 1 and 2). Predicted signal sequence indicated, but may vary by a few positions and depending upon cell type.

35	gcg atg tcg ctc gtg ctg cta agc ctg gcc gcg ctg tgc agg agc gcc	48
	Met Ser Leu Val Leu Leu Ser Leu Ala Ala Leu Cys Arg Ser Ala	
	-10 -5 -1 1	
40	gta ccc cga gag ccg acc gtt caa tgt ggc tct gaa act ggg cca tct	96
	Val Pro Arg Glu Pro Thr Val Gln Cys Gly Ser Glu Thr Gly Pro Ser	
	5 10 15	

	cca gag tgg atg cta caa cat gat cta atc ccg gga gac ttg agg gac	144
	Pro Glu Trp Met Leu Gln His Asp Leu Ile Pro Gly Asp Leu Arg Asp	
	20 25 30	
5	ctc cga gta gaa cct gtt aca act agt gtt gca aca ggg gac tat tca	192
	Leu Arg Val Glu Pro Val Thr Thr Ser Val Ala Thr Gly Asp Tyr Ser	
	35 40 45	
10	att ttg atg aat gta agc tgg gta ctc cgg gca gat gcc agc atc cgc	240
	Ile Leu Met Asn Val Ser Trp Val Leu Arg Ala Asp Ala Ser Ile Arg	
	50 55 60 65	
15	ttg ttg aag gcc acc aag att tgt gtg acg ggc aaa agc aac ttc cag	288
	Leu Leu Lys Ala Thr Lys Ile Cys Val Thr Gly Lys Ser Asn Phe Gln	
	70 75 80	
20	tcc tac agc tgt gtg agg tgc aat tac aca gag gcc ttc cag act cag	336
	Ser Tyr Ser Cys Val Arg Cys Asn Tyr Thr Glu Ala Phe Gln Thr Gln	
	85 90 95	
25	acc aga ccc tct ggt ggt aaa tgg aca ttt tcc tat atc ggc ttc cct	384
	Thr Arg Pro Ser Gly Gly Lys Trp Thr Phe Ser Tyr Ile Gly Phe Pro	
	100 105 110	
30	gta gag ctg aac aca gtc tat ttc att ggg gcc cat aat att cct aat	432
	Val Glu Leu Asn Thr Val Tyr Phe Ile Gly Ala His Asn Ile Pro Asn	
	115 120 125	
35	gca aat atg aat gaa gat ggc cct tcc atg tct gtg aat ttc acc tca	480
	Ala Asn Met Asn Glu Asp Gly Pro Ser Met Ser Val Asn Phe Thr Ser	
	130 135 140 145	
40	cca ggc tgc cta gac cac ata atg aaa tat aaa aaa aag tgt gtc aag	528
	Pro Gly Cys Leu Asp His Ile Met Lys Tyr Lys Lys Lys Cys Val Lys	
	150 155 160	
45	gcc gga agc ctg tgg gat ccg aac atc act gct tgt aag aag aat gag	576
	Ala Gly Ser Leu Trp Asp Pro Asn Ile Thr Ala Cys Lys Lys Asn Glu	
	165 170 175	
50	gag aca gta gaa gtg aac ttc aca acc act ccc ctg gga aac aga tac	624
	Glu Thr Val Glu Val Asn Phe Thr Thr Thr Pro Leu Gly Asn Arg Tyr	
	180 185 190	
55	atg gct ctt atc caa cac agc act atc atc ggg ttt tct cag gtg ttt	672
	Met Ala Leu Ile Gln His Ser Thr Ile Ile Gly Phe Ser Gln Val Phe	
	195 200 205	
60	gag cca cac cag aag aaa caa acg cga gct tca gtg gtg att cca gtg	720
	Glu Pro His Gln Lys Lys Gln Thr Arg Ala Ser Val Val Ile Pro Val	
	210 215 220 225	
65	act ggg gat agt gaa ggt gct acg gtg cag ctg act cca tat ttt cct	768
	Thr Gly Asp Ser Glu Gly Ala Thr Val Gln Leu Thr Pro Tyr Phe Pro	
	230 235 240	

	act tgt ggc agc gac tgc atc cga cat aaa gga aca gtt gtg ctc tgc	816
	Thr Cys Gly Ser Asp Cys Ile Arg His Lys Gly Thr Val Val Leu Cys	
	245 250 255	
5	cca caa aca ggc gtc cct ttc cct ctg gat aac aac aaa agc aag ccg	864
	Pro Gln Thr Gly Val Pro Phe Pro Leu Asp Asn Asn Lys Ser Lys Pro	
	260 265 270	
10	gga ggc tgg ctg cct ctc ctc ctg ctg tct ctg ctg gtg gcc aca tgg	912
	Gly Gly Trp Leu Pro Leu Leu Leu Leu Ser Leu Leu Val Ala Thr Trp	
	275 280 285	
15	gtg ctg gtg gca ggg atc tat cta atg tgg agg cac gaa agg atc aag	960
	Val Leu Val Ala Gly Ile Tyr Leu Met Trp Arg His Glu Arg Ile Lys	
	290 295 300 305	
20	aag act tcc ttt tct acc acc aca cta ctg ccc ccc att aag gtt ctt	1008
	Lys Thr Ser Phe Ser Thr Thr Thr Leu Leu Pro Pro Ile Lys Val Leu	
	310 315 320	
	gtg gtt tac cca tct gaa ata tgt ttc cat cac aca att tgt tac ttc	1056
	Val Val Tyr Pro Ser Glu Ile Cys Phe His His Thr Ile Cys Tyr Phe	
	325 330 335	
25	act gaa ttt ctt caa aac cat tgc aga agt gag gtc atc ctt gaa aag	1104
	Thr Glu Phe Leu Gln Asn His Cys Arg Ser Glu Val Ile Leu Glu Lys	
	340 345 350	
30	tgg cag aaa aag aaa ata gca gag atg ggt cca gtg cag tgg ctt gcc	1152
	Trp Gln Lys Lys Lys Ile Ala Glu Met Gly Pro Val Gln Trp Leu Ala	
	355 360 365	
35	act caa aag aag gca gca gac aaa gtc gtc ttc ctt ctt tcc aat gac	1200
	Thr Gln Lys Lys Ala Ala Asp Lys Val Val Phe Leu Leu Ser Asn Asp	
	370 375 380 385	
40	gtc aac agt gtg tgc gat ggt acc tgt ggc aag agc gag ggc agt ccc	1248
	Val Asn Ser Val Cys Asp Gly Thr Cys Gly Lys Ser Glu Gly Ser Pro	
	390 395 400	
	agt gag aac tct caa gac ctc ttc ccc ctt gcc ttt aac ctt ttc tgc	1296
	Ser Glu Asn Ser Gln Asp Leu Phe Pro Leu Ala Phe Asn Leu Phe Cys	
	405 410 415	
45	agt gat cta aga agc cag att cat ctg cac aaa tac gtg gtg gtc tac	1344
	Ser Asp Leu Arg Ser Gln Ile His Leu His Lys Tyr Val Val Val Tyr	
	420 425 430	
50	ttt aga gag att gat aca aaa gac gat tac aat gct ctc agt gtc tgc	1392
	Phe Arg Glu Ile Asp Thr Lys Asp Asp Tyr Asn Ala Leu Ser Val Cys	
	435 440 445	
55	ccc aag tac cac ctc atg aag gat gcc act gct ttc tgt gca gaa ctt	1440
	Pro Lys Tyr His Leu Met Lys Asp Ala Thr Ala Phe Cys Ala Glu Leu	
	450 455 460 465	



ctc cat gtc aag cag cag gtg tca gca gga aaa aga tca caa gcc tgc 1488  
 Leu His Val Lys Gln Gln Val Ser Ala Gly Lys Arg Ser Gln Ala Cys  
 470 475 480

5 cac gat ggc tgc tgc tcc ttg tagccaccc atgagaagca agagacctta 1539  
 His Asp Gly Cys Cys Ser Leu  
 485

10 aaggcttcct atcccaccaa ttacagggaa aaaacgtgtg atgatcctga agcttactat 1599  
 gcagcctaca aacagcctta gtaattaaaa ctttttatac caataaaatt ttcaaattatt 1659  
 gctaactaat gtagcattaa ctaacgattg gaaactacat ttacaacttc aaagctgttt 1719

15 tatacataga aatcaattac agctttaatt gaaaactgta accattttga taatgcaaca 1779  
 ataaagcatc ttcagcc 1796

20 MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLPGLRDLRVEPVTTSVATGDYSILMNVSWVL  
 RADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIPNA  
 NMNEDGPSMSVNFSTSPGCLDHIMKYKKKCVKAGSLWDPNITACKKNEETVEVNFNTTTPLGNYMALIQHSTI  
 IGFSQVFEPHQKKQTRASVVIPTGDSGATVQLTPYFPPTCGSDCIRHKGTVVLCPTGTGVPFPLDNNKSKPG  
 GWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTTLLPPIKVLVVPSEICFHHTICYFTEFLQNHCR  
 SEVILEKWQKKKIAEMGPVQWLATQKKAADKVFLLSNDVNSVCDGTGCGKSEGSPESENSQDLFPLAFNLFCS  
 25 DLRSQIHLHKYVVVYFREIDTKDDYNALSVC PKYHLMKDATAFCAELLHV KQVVSAGKRSQACHDGCCSL.

Reverse translation of primate, e.g., human, DCRS6 (SEQ ID NO: 3):

30 atgwsnytnng tnytnytnws nytnngcngcn ytntgymgnw sngcngtncc nmngnarcen 60  
 acngtncart gyggwnwsnga racngngccn wsncngart ggatgytnca rcaygayytn 120

35 athccnggng ayytnmnga yytnmngtn garccngtna cnacnwsngt ngcnacnggn 180  
 gaytaywsna thytnatgaa ygtwnsntgg gtnytnmngng cngaygcnws nathmngnytn 240  
 ytnaargcna cnaarathtg ygtnacnggn aarwsnaayt tycarwsnta ywsntgygtn 300

40 mgntgyaayt ayacngargc nttycaracn caracnmngc cnwsnggngg naartggacn 360  
 ttywsntaya thggnttycc ngtngarytn aayacngtnt ayttiyathgg ngcncaayaay 420  
 athccnaayg cnaayatgaa ygargayggn ccnwsnatgw sngtnaaytt yacnwsnccn 480

45 ggntgyytng aycayathat gaartayaar aaraartgyg tnaargcngg nwsnytnngg 540  
 gayccnaaya thacngcntg yaaraaraay gargaracng tngargtnaa yttyacnacn 600

50 acncnytnng gnaaymgnta yatggcnytn athcarcayw snacnathat hggnttywsn 660  
 cargtnnttyg arccncayca raaraarcac acnmngngcnw sngtnngtnat hccngtnacn 720

55 ggngaywsng arggngcnac ngtnrcarytn acncntayt tyccnacntg yggnwsngay 780  
 tgyathmgnc ayaarggnac ngtnngtnytn tgyccncara cngngngtncc nttyccnytn 840  
 gayaayaaya arwsnaarcc ngngngntgg ytnccnytny tnytnytnws nytnytnngtn 900

gcnacntggg tnytngtngc nggnathtay ytnatgtggm gncaygarmg nathaaraar 960  
 acnwsnttyw snacnacnac nytnytncn ccnathaarg tnytngtngt ntayccnwsn 1020  
 5 garathtgyt tycaycayac nathtgytay ttyacngart tyytncaraa ycaytgymgn 1080  
 wsgargtna thytngaraa rtggcaraar aaraarathg cngaratggg nccngtncar 1140  
 10 tggytngcna cncaraaraa rgcngcngay aargtngtnt tyytnytnws naaygaygtn 1200  
 aaywsngtnt gygayggnac ntgyggnaar wsgarggnw snccnwsnga raaywsncar 1260  
 gayytnttyc cnytngcntt yaayytntty tgywsngayy tnmgnwsnca rathcayytn 1320  
 15 cayaartayg tngtngtnta yttymngar athgayacna argaygayta yaaygcnytn 1380  
 wsngtntgyc cnaartayca yytnatgaar gaygcnacng cnttytgygc ngarytnytn 1440  
 20 caygtnaarc arcargtnws ngcnggnaar mgnwsncarg cntgycayga yggntgytgy 1500  
 wsnytn 1506

25 Rodent, e.g., mouse embodiment (see SEQ ID NO: 4 and 5).

gat ttc agc agc cag acg cat ctg cac aaa tac ctg gag gtc tat ctt 48  
 Asp Phe Ser Ser Gln Thr His Leu His Lys Tyr Leu Glu Val Tyr Leu  
 1 5 10 15  
 30 ggg gga gca gac ctc aaa ggc gac tat aat gcc ctg agt gtc tgc ccg 96  
 Gly Gly Ala Asp Leu Lys Gly Asp Tyr Asn Ala Leu Ser Val Cys Pro  
 20 25 30  
 35 caa tat cat ctc atg aag gac gcc aca gct ttc cac aca gaa ctt ctc 144  
 Gln Tyr His Leu Met Lys Asp Ala Thr Ala Phe His Thr Glu Leu Leu  
 35 40 45  
 40 aag gct acg cag agc atg tca gtg aag aaa cgc tca caa gcc tgc cat 192  
 Lys Ala Thr Gln Ser Met Ser Val Lys Lys Arg Ser Gln Ala Cys His  
 50 55 60  
 gat agc tgt tca ccc ttg tagtccaccc gggggaatag agactctgaa 240  
 Asp Ser Cys Ser Pro Leu  
 45 65 70  
 gccttcctac tctcccttcc agtgacaaat gctgtgtgac gactctgaaa tgtgtgggag 300  
 aggctgtgtg gaggtagtgc tatgtacaaa cttgctttaa aactggagtt tgcaaagtca 360  
 50 acctgagcat acacgcctga ggctagtcac tggctggatt tatgaagaca acacagttac 420  
 agacaataat gagtgggacc tacatttggg atatacccaa agctgggtaa tgattatcac 480  
 55 tgagaaccac gcaactctggc catgaggtaa tacggcactt ccctgtcagg ctgtctgtca 540  
 gggtgggtct gtcttgcaact gccatgctc tatgctgcac gtagaccgtt ttgtaacatt 600  
 ttaatctgtt aatgaataat ccgtttggga ggctctc 637

DFSSQTHLHKYLEVYLGGADLKGDNALSVCPQYHLMKDATAFHTELLKATQSMSVKKRSQACHDSCSPL.

5 Reverse translation of rodent, e.g., mouse, DCRS6 (SEQ ID NO: 6):

gayttywsnw sncaracnca yytncayaar tayytnngarg tntayytnng ngngngcngay 60  
 ytnaarggng aytayaaygc nytnwsngtn tgyccncart aycayytnat gaargaygcn 120  
 10 acngcnttyc ayacngaryt nytnaargcn acncarwsna tgwsngtnaa raarmgnwsn 180  
 cargcntgyc aygaywsntg ywsnccnytn 210

15

Table 2: Nucleotide and polypeptide sequences of DNAX Cytokine Receptor Subunit like  
 embodiments (DCRS7). Primate, e.g., human, embodiment (see SEQ ID NO: 7 and 8).  
 Predicted signal sequence indicated, but may vary by a few positions and depending upon cell  
 type.

20

gagtcaggac tcccaggaca gagagtgcac aaactaccca gcacagcccc ctccgcccc 60

tctggaggct gaagagggat tccagcccct gccaccaca gacacgggct gactgggggtg 120

25

tctgcccccc ttgggggcan ccacagggcc tcaggcctgg gtgccacctg gcactagaag 180

atg cct gtg ccc tgg ttc ttg ctg tcc ttg gca ctg ggc cga agc cag 228  
 Met Pro Val Pro Trp Phe Leu Leu Ser Leu Ala Leu Gly Arg Ser Gln  
 -20 -15 -10 -5

30

tgg atc ctt tct ctg gag agg ctt gtg ggg cct cag gac gct acc cac 276  
 Trp Ile Leu Ser Leu Glu Arg Leu Val Gly Pro Gln Asp Ala Thr His  
 -1 1 5 10

35

tgc tct ccg ggc ctc tcc tgc cgc ctc tgg gac agt gac ata ctc tgc 324  
 Cys Ser Pro Gly Leu Ser Cys Arg Leu Trp Asp Ser Asp Ile Leu Cys  
 15 20 25

40

ctg cct ggg gac atc gtg cct gct ccg ggc ccc gtg ctg gcg cct acg 372  
 Leu Pro Gly Asp Ile Val Pro Ala Pro Gly Pro Val Leu Ala Pro Thr  
 30 35 40

45

cac ctg cag aca gag ctg gtg ctg agg tgc cag aag gag acc gac tgt 420  
 His Leu Gln Thr Glu Leu Val Leu Arg Cys Gln Lys Glu Thr Asp Cys  
 45 50 55 60

50

gac ctc tgt ctg cgt gtg gct gtc cac ttg gcc gtg cat ggg cac tgg 468  
 Asp Leu Cys Leu Arg Val Ala Val His Leu Ala Val His Gly His Trp  
 65 70 75

gaa gag cct gaa gat gag gaa aag ttt gga gga gca gct gac tta ggg 516  
 Glu Glu Pro Glu Asp Glu Glu Lys Phe Gly Gly Ala Ala Asp Leu Gly  
 80 85 90

55

gtg gag gag cct agg aat gcc tct ctc cag gcc caa gtc gtg ctc tcc 564  
 Val Glu Glu Pro Arg Asn Ala Ser Leu Gln Ala Gln Val Val Leu Ser  
 95 100 105

	ttc cag gcc tac cct act gcc cgc tgc gtc ctg ctg gag gtg caa gtg	612
	Phe Gln Ala Tyr Pro Thr Ala Arg Cys Val Leu Leu Glu Val Gln Val	
	110 115 120	
5	cct gct gcc ctt gtg cag ttt ggt cag tct gtg ggc tct gtg gta tat	660
	Pro Ala Ala Leu Val Gln Phe Gly Gln Ser Val Gly Ser Val Val Tyr	
	125 130 135 140	
10	gac tgc ttc gag gct gcc cta ggg agt gag gta cga atc tgg tcc tat	708
	Asp Cys Phe Glu Ala Ala Leu Gly Ser Glu Val Arg Ile Trp Ser Tyr	
	145 150 155	
15	act cag ccc agg tac gag aag gaa ctc aac cac aca cag cag ctg cct	756
	Thr Gln Pro Arg Tyr Glu Lys Glu Leu Asn His Thr Gln Gln Leu Pro	
	160 165 170	
20	gac tgc agg ggg ctc gaa gtc tgg aac agc atc ccg agc tgc tgg gcc	804
	Asp Cys Arg Gly Leu Glu Val Trp Asn Ser Ile Pro Ser Cys Trp Ala	
	175 180 185	
25	ctg ccc tgg ctc aac gtg tca gca gat ggt gac aac gtg cat ctg gtt	852
	Leu Pro Trp Leu Asn Val Ser Ala Asp Gly Asp Asn Val His Leu Val	
	190 195 200	
30	ctg aat gtc tct gag gag cag cac ttc ggc ctc tcc ctg tac tgg aat	900
	Leu Asn Val Ser Glu Glu Gln His Phe Gly Leu Ser Leu Tyr Trp Asn	
	205 210 215 220	
35	cag gtc cag ggc ccc cca aaa ccc cgg tgg cac aaa aac ctg act gga	948
	Gln Val Gln Gly Pro Pro Lys Pro Arg Trp His Lys Asn Leu Thr Gly	
	225 230 235	
40	ccg cag atc att acc ttg aac cac aca gac ctg gtt ccc tgc ctc tgt	996
	Pro Gln Ile Ile Thr Leu Asn His Thr Asp Leu Val Pro Cys Leu Cys	
	240 245 250	
45	att cag gtg tgg cct ctg gaa cct gac tcc gtt agg acg aac atc tgc	1044
	Ile Gln Val Trp Pro Leu Glu Pro Asp Ser Val Arg Thr Asn Ile Cys	
	255 260 265	
50	ccc ttc agg gag gac ccc cgc gca cac cag aac ctc tgg caa gcc gcc	1092
	Pro Phe Arg Glu Asp Pro Arg Ala His Gln Asn Leu Trp Gln Ala Ala	
	270 275 280	
55	cga ctg cga ctg ctg acc ctg cag agc tgg ctg ctg gac gca ccg tgc	1140
	Arg Leu Arg Leu Leu Thr Leu Gln Ser Trp Leu Leu Asp Ala Pro Cys	
	285 290 295 300	
55	tcg ctg ccc gca gaa gcg gca ctg tgc tgg cgg gct ccg ggt ggg gac	1188
	Ser Leu Pro Ala Glu Ala Ala Leu Cys Trp Arg Ala Pro Gly Gly Asp	
	305 310 315	
55	ccc tgc cag cca ctg gtc cca ccg ctt tcc tgg gag aat gtc act gtg	1236
	Pro Cys Gln Pro Leu Val Pro Pro Leu Ser Trp Glu Asn Val Thr Val	
	320 325 330	
55	gac gtg aac agc tcg gag aag ctg cag ctg cag gag tgc ttg tgg gct	1284
	Asp Val Asn Ser Ser Glu Lys Leu Gln Leu Gln Glu Cys Leu Trp Ala	
	335 340 345	

5	gac tcc ctg ggg cct ctc aaa gac gat gtg cta ctg ttg gag aca cga Asp Ser Leu Gly Pro Leu Lys Asp Asp Val Leu Leu Leu Glu Thr Arg 350 355 360	1332
10	ggc ccc cag gac aac aga tcc ctc tgt gcc ttg gaa ccc agt ggc tgt Gly Pro Gln Asp Asn Arg Ser Leu Cys Ala Leu Glu Pro Ser Gly Cys 365 370 375 380	1380
15	act tca cta ccc agc aaa gcc tcc acg agg gca gct cgc ctt gga gag Thr Ser Leu Pro Ser Lys Ala Ser Thr Arg Ala Ala Arg Leu Gly Glu 385 390 395	1428
20	tac tta cta caa gac ctg cag tca ggc cag tgt ctg cag cta tgg gac Tyr Leu Leu Gln Asp Leu Gln Ser Gly Gln Cys Leu Gln Leu Trp Asp 400 405 410	1476
25	gat gac ttg gga gcg cta tgg gcc tgc ccc atg gac aaa tac atc cac Asp Asp Leu Gly Ala Leu Trp Ala Cys Pro Met Asp Lys Tyr Ile His 415 420 425	1524
30	aag cgc tgg gcc ctc gtg tgg ctg gcc tgc cta ctc ttt gcc gct gcg Lys Arg Trp Ala Leu Val Trp Leu Ala Cys Leu Leu Phe Ala Ala Ala 430 435 440	1572
35	ctt tcc ctc atc ctc ctt ctc aaa aag gat cac gcg aaa ggg tgg ctg Leu Ser Leu Ile Leu Leu Leu Lys Lys Asp His Ala Lys Gly Trp Leu 445 450 455 460	1620
40	agg ctc ttg aaa cag gac gtc cgc tcg ggg gcg gcc gcc agg ggc cgc Arg Leu Leu Lys Gln Asp Val Arg Ser Gly Ala Ala Ala Arg Gly Arg 465 470 475	1668
45	gcg gct ctg ctc ctc tac tca gcc gat gac tcg ggt ttc gag cgc ctg Ala Ala Leu Leu Leu Tyr Ser Ala Asp Asp Ser Gly Phe Glu Arg Leu 480 485 490	1716
50	gtg ggc gcc ctg gcg tcg gcc ctg tgc cag ctg ccg ctg cgc gtg gcc Val Gly Ala Leu Ala Ser Ala Leu Cys Gln Leu Pro Leu Arg Val Ala 495 500 505	1764
55	gta gac ctg tgg agc cgt cgt gaa ctg agc gcg cag ggg ccc gtg gct Val Asp Leu Trp Ser Arg Arg Glu Leu Ser Ala Gln Gly Pro Val Ala 510 515 520	1812
60	tgg ttt cac gcg cag cgg cgc cag acc ctg cag gag ggc ggc gtg gtg Trp Phe His Ala Gln Arg Arg Gln Thr Leu Gln Glu Gly Gly Val Val 525 530 535 540	1860
65	gtc ttg ctc ttc tct ccc ggt gcg gtg gcg ctg tgc agc gag tgg cta Val Leu Leu Phe Ser Pro Gly Ala Val Ala Leu Cys Ser Glu Trp Leu 545 550 555	1908
70	cag gat ggg gtg tcc ggg ccc ggg gcg cac ggc ccg cac gac gcc ttc Gln Asp Gly Val Ser Gly Pro Gly Ala His Gly Pro His Asp Ala Phe 560 565 570	1956

	ccc gcc tcg ctc agc tgc gtg ctg ccc gac ttc ttg cag ggc cgg gcg	2004
	Arg Ala Ser Leu Ser Cys Val Leu Pro Asp Phe Leu Gln Gly Arg Ala	
	575 580 585	
5	ccc gcc agc tac gtg ggg gcc tgc ttc gac agg ctg ctc cac ccg gac	2052
	Pro Gly Ser Tyr Val Gly Ala Cys Phe Asp Arg Leu Leu His Pro Asp	
	590 595 600	
10	gcc gta ccc gcc ctt ttc cgc acc gtg ccc gtc ttc aca ctg ccc tcc	2100
	Ala Val Pro Ala Leu Phe Arg Thr Val Pro Val Phe Thr Leu Pro Ser	
	605 610 615 620	
15	caa ctg cca gac ttc ctg ggg gcc ctg cag cag cct cgc gcc ccg cgt	2148
	Gln Leu Pro Asp Phe Leu Gly Ala Leu Gln Gln Pro Arg Ala Pro Arg	
	625 630 635	
20	tcc ggg cgg ctc caa gag aga gcg gag caa gtg tcc cgg gcc ctt cag	2196
	Ser Gly Arg Leu Gln Glu Arg Ala Glu Gln Val Ser Arg Ala Leu Gln	
	640 645 650	
25	cca gcc ctg gat agc tac ttc cat ccc ccg ggg acn tcc gcg ccg gga	2244
	Pro Ala Leu Asp Ser Tyr Phe His Pro Pro Gly Xaa Ser Ala Pro Gly	
	655 660 665	
30	ccc ggg gtg gga cca ggg gcg gga cct ggg gcg ggg gac ggg act	2289
	Arg Gly Val Gly Pro Gly Ala Gly Pro Gly Ala Gly Asp Gly Thr	
	670 675 680	
35	ttaaataaagg cagacgctg	2308
40	MPVPWFLLSLALGRSQWILSLERLVGPQDATHCSPGLSCLWSDILCLPGDIVPAPGPVLAPTHLQTELVL RCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADLGVEEPRNASLQAQVVLVSFQAYPTARCVLLEVQV PAALVQFGQSVGSVVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPDCRGLEVWNSIPSCWALPWLNVSA DGDNVHLVLNVSEEQHFGLSLYWNQVQGPCKPRWHKNLTGPGQIITLNHTDLVPCLCIQVWPLEPDSVRTNIC PFREDPRAHQNLWQAARLRLTLQSWLLDAPCSLPAEAALCWRAAGGDPQCPLVPPLSWENVTVDVNSSEKL QLQECWLADSLGPLKDDVLLLETRGPQDNRLCALEPSGCTSLPSKASTRAARLGEYLLQDLQSGQCLQLWD DDLGALWACPMCKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSGAAARGRAALLLYSA DDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAVFWFAQRRLQEGGVVLLFSPGAVALCSEWL QDGVSGPGAGHPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLHLPDAVPALFRTVPVFTLPSQLPDLFLGA LQQPRAPRSGRLQERAEQVSRALQPALDSYFHPGTSAPGRGVGPGAGPGAGDGT.	

Reverse translation of primate, e.g., human, DCRS7 (SEQ ID NO: 9):

45	atgccngtnc cntgggttyt nytnwsnytn gcnytnngnm gnwsncartg gathytnwsn	60
	ytngarmgny tngtnggncc ncargaygcn acncaytgyw snccnggnyt nwsntgymgn	120
50	ytntgggayw sngayathyt ntgyytnccn ggngayathg tncngcncc nggncngtn	180
	ytngcnccna cncayytnc racngarytn gtnytnmgnt gycaraarga racngaytgy	240
	gayytnntgyy tnmngntngc ngtncayytn gcngtncayg gncaytggga rgarccngar	300
55	gaygargara arttyggngg ngcngcngay ytngngtng argarccnmg naaygcnwsn	360
	ytncargcnc argtngtnyt nwsnttycar gentayccna cngcnmgntg ygtnytnytn	420
	gargtnarg tncngcngc nytngtncar ttyggncarw sngtnggnws ngtngtntay	480

5      gaytgyttyg argcngcnyt nggnwsngar gtnmgnatht ggwsntayac ncarccnmgn 540  
 taygaraarg arytnaayca yacncarcar ytncngayt gymngggnyt ngargtntgg 600  
 aaywsnathc cnwsntgytg ggcnytnccn tggytnaayg tnwsngcnga yggngayaay 660  
 gtncayytng tnytnaaygt nwsngargar carcayttyg gnytnwsnyt ntaytggay 720  
 10      cargtncarg gncncnca rccnmgttg cayaaraay tncnggnc ncarathath 780  
 acnytnaayc ayacngayt ngtnccntgy ytntgyathc argtntggcc nytngarccn 840  
 15      gaywsngtnm gnacnaayat htgyccntty mgngargayc cnmgngcnc ycaraaytn 900  
 tggcargcng cnmgnytnmg nytnytnacn ytncarwsnt ggytnytnga ygcncntgy 960  
 wsnytnccng cngargcngc nytntgytg mgngcncng gngngaycc ntgyarccn 1020  
 20      ytngtnccnc cnytnwsntg ggaraaygt acngtngayg tnaaywsnws ngaraarytn 1080  
 carytncarg artgyytntg ggcngaywsn ytnggncny tnaargayga ygtnytnyt 1140  
 ytngaracnm gnggncnca rgayaaymgn wsnyntgyg cnytngarcc nwsnggntgy 1200  
 25      acnwsnytn cwnsnaargc nwsnacnmgn gcngcnmgy tngngarta yytnytncar 1260  
 gayytncarw snggncartg yytnarytn tgggaygayg ayytnggngc nytntgggcn 1320  
 30      tgyccnatgg ayaartayat hcayaarmgn tgggcnytn tntggytngc ntgyytnyt 1380  
 ttygcngcng cnytnwsnyt nathytnyt ytnaaraarg aycaygcnaa rggntggytn 1440  
 35      mgnytnyt na arcargaygt nmgnwsnggn gcngcngcnm gngnmngnc ngcnytnyt 1500  
 ytntaywsng cngaygayws nggnttygar mgnytngtng gngcnytnge nwsngcnytn 1560  
 tgycarytn cnytnmgnt ngcngtngay ytntggwsnm gnmngaryt nwsngcncar 1620  
 40      ggnccngtng cntggtyca ygcncarmgn mgncaracny tncargarg gngngtngtn 1680  
 gtnytnytnt tywsncngg ngcngtngcn ytntgywsng artggytnca rgayggngtn 1740  
 45      wsnggncng gngncaygg nccncaygay gcnttymng cwnsnytnws ntgygtnyt 1800  
 ccngaytty tncarggnmg ngcncnggn wsntaygtng gngcntgyt ygaymnytn 1860  
 ytncaycng aygngtncc ngcnytnntt mgnacngtn cngtnntyac nytncnwsn 1920  
 50      carytnccng aytyytngg ngcnytnear carccnmng cncnmgnws nggnmgnytn 1980  
 cargarmng cngarcargt nwsnmngcn ytncarccng cnytngayws ntayttycay 2040  
 55      ccncnggna cwnsgncnc nggnmgnggn gtnggncng gngcnggnc ngngcnggn 2100  
 gayggnacn 2109

Rodent, e.g., mouse, embodiment (see SEQ ID NO: 10 and 11). Predicted signal sequence indicated, but may vary by a few positions and depending upon cell type.

5	ccaaatcgaa agcacgggag ctgatactgg gcctggagtc caggctcact ggagtgggga	60
	agcatggctg gagaggaatt ctagcccttg ctctctccca gggacacggg gctgattgtc	120
	agcaggggag aggggtctgc ccccccttgg gggggcagga cggggcctca ggcctgggtg	180
10	ctgtccggca cctggaag atg cct gtg tcc tgg ttc ctg ctg tcc ttg gca	231
	Met Pro Val Ser Trp Phe Leu Leu Ser Leu Ala	
	-20 -15 -10	
15	ctg ggc cga aac cct gtg gtc gtc tct ctg gag aga ctg atg gag cct	279
	Leu Gly Arg Asn Pro Val Val Val Ser Leu Glu Arg Leu Met Glu Pro	
	-5 -1 1 5	
20	cag gac act gca cgc tgc tct cta ggc ctc tcc tgc cac ctc tgg gat	327
	Gln Asp Thr Ala Arg Cys Ser Leu Gly Leu Ser Cys His Leu Trp Asp	
	10 15 20	
25	ggt gac gtg ctc tgc ctg cct gga agc ctc cag tct gcc cca ggc cct	375
	Gly Asp Val Leu Cys Leu Pro Gly Ser Leu Gln Ser Ala Pro Gly Pro	
	25 30 35	
30	gtg cta gtg cct acc cgc ctg cag acg gag ctg gtg ctg agg tgt cca	423
	Val Leu Val Pro Thr Arg Leu Gln Thr Glu Leu Val Leu Arg Cys Pro	
	40 45 50 55	
35	cag aag aca gat tgc gcc ctc tgt gtc cgt gtg gtg gtc cac ttg gcc	471
	Gln Lys Thr Asp Cys Ala Leu Cys Val Arg Val Val Val His Leu Ala	
	60 65 70	
40	gtg cat ggg cac tgg gca gag cct gaa gaa gct gga aag tct gat tca	519
	Val His Gly His Trp Ala Glu Pro Glu Glu Ala Gly Lys Ser Asp Ser	
	75 80 85	
45	gaa ctc cag gag tct agg aac gcc tct ctc cag gcc cag gtg gtg ctc	567
	Glu Leu Gln Glu Ser Arg Asn Ala Ser Leu Gln Ala Gln Val Val Leu	
	90 95 100	
50	tcc ttc cag gcc tac ccc atc gcc cgc tgt gcc ctg ctg gag gtc cag	615
	Ser Phe Gln Ala Tyr Pro Ile Ala Arg Cys Ala Leu Leu Glu Val Gln	
	105 110 115	
55	gtg ccc gct gac ctg gtg cag cct ggt cag tcc gtg ggt tct gcg gta	663
	Val Pro Ala Asp Leu Val Gln Pro Gly Gln Ser Val Gly Ser Ala Val	
	120 125 130 135	
50	ttt gac tgt ttc gag gct agt ctt ggg gct gag gta cag atc tgg tcc	711
	Phe Asp Cys Phe Glu Ala Ser Leu Gly Ala Glu Val Gln Ile Trp Ser	
	140 145 150	
55	tac acg aag ccc agg tac cag aaa gag ctc aac ctc aca cag cag ctg	759
	Tyr Thr Lys Pro Arg Tyr Gln Lys Glu Leu Asn Leu Thr Gln Gln Leu	
	155 160 165	



	cct gac tgc agg ggt ctt gaa gtc cgg gac agc atc cag agc tgc tgg	807
	Pro Asp Cys Arg Gly Leu Glu Val Arg Asp Ser Ile Gln Ser Cys Trp	
	170 175 180	
5	gtc ctg ccc tgg ctc aat gtg tct aca gat ggt gac aat gtc ctt ctg	855
	Val Leu Pro Trp Leu Asn Val Ser Thr Asp Gly Asp Asn Val Leu Leu	
	185 190 195	
10	aca ctg gat gtc tct gag gag cag gac ttt agc ttc tta ctg tac ctg	903
	Thr Leu Asp Val Ser Glu Glu Gln Asp Phe Ser Phe Leu Leu Tyr Leu	
	200 205 210 215	
15	cgt cca gtc ccg gat gct ctc aaa tcc ttg tgg tac aaa aac ctg act	951
	Arg Pro Val Pro Asp Ala Leu Lys Ser Leu Trp Tyr Lys Asn Leu Thr	
	220 225 230	
20	gga cct cag aac att act tta aac cac aca gac ctg gtt ccc tgc ctc	999
	Gly Pro Gln Asn Ile Thr Leu Asn His Thr Asp Leu Val Pro Cys Leu	
	235 240 245	
25	tgc att cag gtg tgg tcg cta gag cca gac tct gag agg gtc gaa ttc	1047
	Cys Ile Gln Val Trp Ser Leu Glu Pro Asp Ser Glu Arg Val Glu Phe	
	250 255 260	
30	tgc ccc ttc cgg gaa gat ccc ggt gca cac agg aac ctc tgg cac ata	1095
	Cys Pro Phe Arg Glu Asp Pro Gly Ala His Arg Asn Leu Trp His Ile	
	265 270 275	
35	gcc agg ctg cgg gta ctg tcc cca ggg gta tgg cag cta gat gcg cct	1143
	Ala Arg Leu Arg Val Leu Ser Pro Gly Val Trp Gln Leu Asp Ala Pro	
	280 285 290 295	
40	tgc tgt ctg ccg ggc aag gta aca ctg tgc tgg cag gca cca gac cag	1191
	Cys Cys Leu Pro Gly Lys Val Thr Leu Cys Trp Gln Ala Pro Asp Gln	
	300 305 310	
45	agt ccc tgc cag cca ctt gtg cca cca gtg ccc cag aag aac gcc act	1239
	Ser Pro Cys Gln Pro Leu Val Pro Pro Val Pro Gln Lys Asn Ala Thr	
	315 320 325	
50	gtg aat gag cca caa gat ttc cag ttg gtg gca ggc cac ccc aac ctc	1287
	Val Asn Glu Pro Gln Asp Phe Gln Leu Val Ala Gly His Pro Asn Leu	
	330 335 340	
55	tgt gtc cag gtg agc acc tgg gag aag gtt cag ctg caa gcg tgc ttg	1335
	Cys Val Gln Val Ser Thr Trp Glu Lys Val Gln Leu Gln Ala Cys Leu	
	345 350 355	
60	tgg gct gac tcc ttg ggg ccc ttc aag gat gat atg ctg tta gtg gag	1383
	Trp Ala Asp Ser Leu Gly Pro Phe Lys Asp Asp Met Leu Leu Val Glu	
	360 365 370 375	
65	atg aaa acc ggc ctc aac aac aca tca gtc tgt gcc ttg gaa ccc agt	1431
	Met Lys Thr Gly Leu Asn Asn Thr Ser Val Cys Ala Leu Glu Pro Ser	
	380 385 390	
70	ggc tgt aca cca ctg ccc agc atg gcc tcc acg aga gct gct cgc ctg	1479
	Gly Cys Thr Pro Leu Pro Ser Met Ala Ser Thr Arg Ala Ala Arg Leu	
	395 400 405	

5	gga gag gag ttg ctg caa gac ttc cga tca cac cag tgt atg cag ctg																1527
	Gly	Glu	Glu	Leu	Leu	Gln	Asp	Phe	Arg	Ser	His	Gln	Cys	Met	Gln	Leu	
	410						415						420				
10	tgg aac gat gac aac atg gga tgc cta tgg gcc tgc ccc atg gac aag																1575
	Trp	Asn	Asp	Asp	Asn	Met	Gly	Ser	Leu	Trp	Ala	Cys	Pro	Met	Asp	Lys	
	425						430						435				
15	tac atc cac agg cgc tgg gtc cta gta tgg ctg gcc tgc cta ctc ttg																1623
	Tyr	Ile	His	Arg	Arg	Trp	Val	Leu	Val	Trp	Leu	Ala	Cys	Leu	Leu	Leu	
	440						445						450			455	
20	gct gcg gcg ctt ttc ttc ttc ctc ctt cta aaa aag gac cgc agg aaa																1671
	Ala	Ala	Ala	Leu	Phe	Phe	Phe	Leu	Leu	Leu	Lys	Lys	Asp	Arg	Arg	Lys	
				460						465						470	
25	gcg gcc cgt ggc tcc cgc acg gcc ttg ctc ctc cac tcc gcc gac gga																1719
	Ala	Ala	Arg	Gly	Ser	Arg	Thr	Ala	Leu	Leu	Leu	His	Ser	Ala	Asp	Gly	
				475						480						485	
30	gcg ggc tac gag cgc ctg gtg gga gca ctg gcg tcc gcg ttg agc cag																1767
	Ala	Gly	Tyr	Glu	Arg	Leu	Val	Gly	Ala	Leu	Ala	Ser	Ala	Leu	Ser	Gln	
				490						495						500	
35	atg cca ctg cgc gtg gcc gtg gac ctg tgg agc cgc cgc gag ctg agc																1815
	Met	Pro	Leu	Arg	Val	Ala	Val	Asp	Leu	Trp	Ser	Arg	Arg	Glu	Leu	Ser	
				505						510						515	
40	gcg cac gga gcc cta gcc tgg ttc cac cac cag cga cgc cgt atc ctg																1863
	Ala	His	Gly	Ala	Leu	Ala	Trp	Phe	His	His	Gln	Arg	Arg	Arg	Ile	Leu	
	520						525						530			535	
45	cag gag ggt ggc gtg gta atc ctt ctc ttc tgc ccc gcg gcc gtg gcg																1911
	Gln	Glu	Gly	Gly	Val	Val	Ile	Leu	Leu	Phe	Ser	Pro	Ala	Ala	Val	Ala	
				540						545						550	
50	cag tgt cag cag tgg ctg cag ctc cag aca gtg gag ccc ggg ccg cat																1959
	Gln	Cys	Gln	Gln	Trp	Leu	Gln	Leu	Gln	Thr	Val	Glu	Pro	Gly	Pro	His	
				555						560						565	
55	gac gcc ctc gcc gcc tgg ctc agc tgc gtg cta ccc gat ttc ctg caa																2007
	Asp	Ala	Leu	Ala	Ala	Trp	Leu	Ser	Cys	Val	Leu	Pro	Asp	Phe	Leu	Gln	
				570						575						580	
60	ggc cgg gcg acc ggc cgc tac gtc ggg gtc tac ttc gac ggg ctg ctg																2055
	Gly	Arg	Ala	Thr	Gly	Arg	Tyr	Val	Gly	Val	Tyr	Phe	Asp	Gly	Leu	Leu	
				585						590						595	
65	cac cca gac tct gtg ccc tcc ccg ttc cgc gtc gcc ccg ctc ttc tcc																2103
	His	Pro	Asp	Ser	Val	Pro	Ser	Pro	Phe	Arg	Val	Ala	Pro	Leu	Phe	Ser	
	600						605						610			615	
70	ctg ccc tgc cag ctg ccg gct ttc ctg gat gca ctg cag gga ggc tgc																2151
	Leu	Pro	Ser	Gln	Leu	Pro	Ala	Phe	Leu	Asp	Ala	Leu	Gln	Gly	Gly	Cys	
				620						625						630	

tcc act tcc gcg ggg cga ccc gcg gac cgg gtg gaa cga gtg acc cag 2199  
 Ser Thr Ser Ala Gly Arg Pro Ala Asp Arg Val Glu Arg Val Thr Gln  
                     635                    640                    645

5     gcg ctg cgg tcc gcc ctg gac agc tgt act tct agc tcg gaa gcc cca 2247  
       Ala Leu Arg Ser Ala Leu Asp Ser Cys Thr Ser Ser Ser Glu Ala Pro  
                     650                    655                    660

10    ggc tgc tgc gag gaa tgg gac ctg gga ccc tgc act aca cta gaa 2292  
       Gly Cys Cys Glu Glu Trp Asp Leu Gly Pro Cys Thr Thr Leu Glu  
                     665                    670                    675

taaaagccga tacagtattc ct 2314

15    MPVSWFLLSLALGRNPVVVSLERLMEPQDTARCSLGLSCHLWDGDVLCPLPGSLQSAPGPVLPVTRLQTELVL  
       RCPQKTDCAVCVRVVHLAVHGHWAEPPEEAGKSDSELQESRNASLQAQVVLSFQAYPIARCALLEVQVPADL  
       VQPGQSVGSAVFDCFEASLGAEVQIWSYTKPRYQKELNLTQQLPDCRGLEVRDSIQSCWVLPWLVNSTDGDN  
       VLLTLDVSEEQDFSFLLYLRPVPDALKSLWYKNLTGPQNITLNTDLVPCLCIQVWSLEPDSEVEFCPFRE  
       DPGAHRLNWHIARLRVLSPGVWQLDAPCCLPGKVTLWCQAPDQSPCQPLVPPVPQKNATVNEPQDFQLVAGH  
       PNLCVQVSTWEKVQLQACLWADSLGPFKDDMLLVEMKTGLNNTSVCALPSGCTPLPSMASTRAARLGEELL  
       QDFRSHQCMQLWDDNMGSLWACPMDKYIHRRWVVLWLACLLAAALFFFLLLKKDRRKAARGSRRTALLLHS  
       ADGAGYERLVGALASALSQMPLRVAVDLWSRRELSAHGALAWFHHQRRRILQEGGVVILLFSPAAVAQCQW  
       LQLQTVEPGPHDALAAWLSCVLPDFLQGRATGRYGVYFDGLLHPDSVPSPFRVAPLPSQLPAFLDALQ  
       GGCSTSAGRPADRVERVTQALRSALDSCTSSSEAPGCCEWDLGPCTTLE.

20

25

Reverse translation of rodent, e.g., mouse, DCRS7 (SEQ ID NO: 12):

atgccngtnw sntggttyt nytnwsnytn gcnytngnm gnaaycngt ngtnngtnwsn 60

30     ytngarmgny tnatggarcc ncargayacn gcnmgntgyw snytnggnyt nwsntgycay 120

      ytntgggayg gngaygtnyt ntgyytnccn ggnwsnytn arwsngcnc nggncngtn 180

35     ytngtnccna cnmgnytnca racngarytn gtnytnmgnt gyccncaraa racngaytgy 240

      gcnytntgyg tnmngntngt ngtncayytn gcngtncayg gncaytgggc ngarccngar 300

      gargcnggna arwsngayws ngarytnear garwsnmgna aygcnwsnyt ncargcncar 360

40     gtngtnytnw snttycargc ntayccnath gcnmgntgyg cnytnytnga rgtncargtn 420

      ccngcngayy tngtnearcc nggncarwsn gtnggnwsng cngtnttyga ytgyttygar 480

45     gcnwsnytn gngcngargt ncarathtgg wsntayacna arccnmngnta ycaraargar 540

      ytnaaaytna cncarcaryt nccngaytgy mgnggnytn argtnmgnga ywsnathcar 600

      wsntgytggg tnytnccntg gytnaaygtn wsnaacngayg gngayaaygt nytnytnacn 660

50     ytngaygtnw sngargarca rgayttywsn tyytnytn tnytnmgnc ngtnccngay 720

      gcnytnaarw snytntggt yaaraaaytn acnggncnc araayathac nytnaaycay 780

55     acngayytn tncntgyt ntgyathcar gtntggwsny tngarccnga ywsngarmgn 840

      gtngarttyt gycenttymg ngargaycnc ggngcncaym gnaaayytn gcaayathgcn 900

      mgnytnmgng tnytnwsncc ngngntntgg carytngayg cncntgyt yytncnggn 960

aargtnacny tntgytggca rgcnccngay carwsnccnt gycarccnyt ngtnccnccn 1020  
 5 gtnccncara araaygcnae ngtnaaygar ccncargayt tycarytngt ngcnggncay 1080  
 ccnaaaytnt gygtncargt nwsnacntgg garaargtnc arytnccargc ntgyytntgg 1140  
 gcngaywsny tnggnccntt yaargaygay atgytntng tngaratgaa racnggnytn 1200  
 10 aayaayacnw sngtntgygc nytnargarccn wsgngntgya cncnytncc nwsnatggcn 1260  
 wsnacnmng cngcnmngyt ngngngargar ytntncarg ayttymgnws ncaycartgy 1320  
 atgcarytnt ggaaygayga yaayatgggn wsnytntggg cntgyccnat ggayaartay 1380  
 15 athcaymngm gntgggtnt ntntgggytn gcntgyytny tntngcngc ngcnytnntt 1440  
 ttyttytny tnytnaaraa rgaymngmgn aargcngcnm gnggnwsnmg nacngcnytn 1500  
 20 ytntncayw sngcngaygg ngcnggntay garmgnytn tngngcnytn ngcnwsngcn 1560  
 ytnwsncara tgccnytnmg ngtnngcngtn gayytnntggw snmngmngna rytnwsngcn 1620  
 cayggngcny tngcntggtt ycaycaycar mgnmngmngna thytncarga rggnggngtn 1680  
 25 gtnathytny tnttywsncc ngcngcngtn gcncartgyc arcartggyt ncarytnear 1740  
 acngtnargc cnggnccnca ygaygcnytn gcngcntggy tnwsntgygt nytncngay 1800  
 30 ttytncarg gnmngcnae ngngmngntay gtngngntnt aytygaygg nytnytnay 1860  
 ccngaywsng tncnwsncc ntymngntn gcncnytn tywsnytncc nwancarytn 1920  
 ccngcnttyy tngaygcnytn ncargggngn tgywsnacnw sngcnggngm nccngcngay 1980  
 35 mngntngarm gngtnacnae rgcnytnmgn wsgcnytn aywsntgyac nwsnwsnwsn 2040  
 gargcncng gntgytgyga rgartgggay ytnggnccnt gyacnacnytn ngar 2094

40 Table 3: Nucleotide and polypeptide sequences of DNAX Cytokine Receptor Subunit like  
 embodiments (DCRS8). Primate, e.g., human, embodiment (see SEQ ID NO: 13 and 14).  
 Predicted signal sequence indicated, but may vary by a few positions and depending upon cell  
 type.

45 cccacgcntc cgggccagca gcggcgcc gggcgccaga gaacggcctg gctgggag 60  
 cgcacggcc atg gcc ccg tgg ctg cag etc tgc tcc gtc ttc ttt acg gtc 111  
 Met Ala Pro Trp Leu Gln Leu Cys Ser Val Phe Phe Thr Val  
 50 -15 -10 -5  
 aac gcc tgc etc aac ggc tcg cag ctg gct gtn gcc gct ggc ggg tcc 159  
 Asn Ala Cys Leu Asn Gly Ser Gln Leu Ala Xaa Ala Ala Gly Gly Ser  
 -1 1 5 10  
 55 ggc cgc gcg cng ggc gcc gac acc tgt agc tgg ang gga gtg ggg cca 207  
 Gly Arg Ala Xaa Gly Ala Asp Thr Cys Ser Trp Xaa Gly Val Gly Pro  
 15 20 25 30

5	gcc agc aga aac agt ggg ctg tac aac atc acc ttc aaa tat gac aat 255 Ala Ser Arg Asn Ser Gly Leu Tyr Asn Ile Thr Phe Lys Tyr Asp Asn	35 40 45
	tgt acc acc tac ttg aat cca gtg ggg aag cat gtg att gct gac gcc 303 Cys Thr Thr Tyr Leu Asn Pro Val Gly Lys His Val Ile Ala Asp Ala	50 55 60
10	cag aat atc acc atc agc cag tat gct tgc cat gac caa gtg gca gtc 351 Gln Asn Ile Thr Ile Ser Gln Tyr Ala Cys His Asp Gln Val Ala Val	65 70 75
	acc att ctt tgg tcc cca ggg gcc ctc ggc atc gaa ttc ctg aaa gga 399 Thr Ile Leu Trp Ser Pro Gly Ala Leu Gly Ile Glu Phe Leu Lys Gly	80 85 90
20	ttt cgg gta ata ctg gag gag ctg aag tcg gag gga aga cag ngc caa 447 Phe Arg Val Ile Leu Glu Glu Leu Lys Ser Glu Gly Arg Gln Xaa Gln	95 100 105 110
	caa ctg att cta aag gat ccg aag cag ntc aac agt agc ttc aaa aga 495 Gln Leu Ile Leu Lys Asp Pro Lys Gln Xaa Asn Ser Ser Phe Lys Arg	115 120 125
25	act gga atg gaa tct caa cct ttn ctg aat atg aaa ttt gaa acg gat 543 Thr Gly Met Glu Ser Gln Pro Xaa Leu Asn Met Lys Phe Glu Thr Asp	130 135 140
	tat ttc gta agg ttg tcc ttt tcc ttc att aaa aac gaa agc aat tac 591 Tyr Phe Val Arg Leu Ser Phe Ser Phe Ile Lys Asn Glu Ser Asn Tyr	145 150 155
35	cac cct ttc ttc ttt aga acc cga gcc tgt gac ctg ttg tta cag ccg 639 His Pro Phe Phe Phe Arg Thr Arg Ala Cys Asp Leu Leu Leu Gln Pro	160 165 170
	gac aat cta gct tgt aaa ccc ttc tgg aag cct cgg aac ctg aac atc 687 Asp Asn Leu Ala Cys Lys Pro Phe Trp Lys Pro Arg Asn Leu Asn Ile	175 180 185 190
40	agc cag cat ggc tcg gac atg cag gtg tcc ttc gac cac gca ccg cac 735 Ser Gln His Gly Ser Asp Met Gln Val Ser Phe Asp His Ala Pro His	195 200 205
	aac ttc ggc ttc cgt ttc ttc tat ctt cac tac aag ctc aag cac gaa 783 Asn Phe Gly Phe Arg Phe Phe Tyr Leu His Tyr Lys Leu Lys His Glu	210 215 220
50	gga cct ttc aag cga aag acc tgt aag cag gag caa act aca gag atg 831 Gly Pro Phe Lys Arg Lys Thr Cys Lys Gln Glu Gln Thr Thr Glu Met	225 230 235
	acc agc tgc ctc ctt caa aat gtt tct cca ggg gat tat ata att gag 879 Thr Ser Cys Leu Leu Gln Asn Val Ser Pro Gly Asp Tyr Ile Ile Glu	240 245 250

	ctg	gtg	gat	gac	act	aac	aca	aca	aga	aaa	gtg	atg	cat	tat	gcc	tta	927
	Leu	Val	Asp	Asp	Thr	Asn	Thr	Thr	Arg	Lys	Val	Met	His	Tyr	Ala	Leu	
	255					260					265					270	
5	aag	cca	gtg	cac	tcc	ccg	tgg	gcc	ggg	ccc	atc	aga	gcc	gtg	gcc	atc	975
	Lys	Pro	Val	His	Ser	Pro	Trp	Ala	Gly	Pro	Ile	Arg	Ala	Val	Ala	Ile	
					275					280					285		
10	aca	gtg	cca	ctg	gta	gtc	ata	tcg	gca	ttc	gcg	acg	ctc	ttc	act	gtg	1023
	Thr	Val	Pro	Leu	Val	Val	Ile	Ser	Ala	Phe	Ala	Thr	Leu	Phe	Thr	Val	
				290					295					300			
15	atg	tgc	cgc	aag	aag	caa	caa	gaa	aat	ata	tat	tca	cat	tta	gat	gaa	1071
	Met	Cys	Arg	Lys	Lys	Gln	Gln	Glu	Asn	Ile	Tyr	Ser	His	Leu	Asp	Glu	
			305					310					315				
20	gag	agc	tct	gag	tct	tcc	aca	tac	act	gca	gca	ctc	cca	aga	gag	agg	1119
	Glu	Ser	Ser	Glu	Ser	Ser	Thr	Tyr	Thr	Ala	Ala	Leu	Pro	Arg	Glu	Arg	
		320					325					330					
	ctc	cgg	ccg	cgg	ccg	aag	gtc	ttt	ctc	tgc	tat	tcc	agt	aaa	gat	ggc	1167
	Leu	Arg	Pro	Arg	Pro	Lys	Val	Phe	Leu	Cys	Tyr	Ser	Ser	Lys	Asp	Gly	
	335					340					345				350		
25	cag	aat	cac	atg	aat	gtc	gtc	cag	tgt	ttc	gcc	tac	ttc	ctc	cag	gac	1215
	Gln	Asn	His	Met	Asn	Val	Val	Gln	Cys	Phe	Ala	Tyr	Phe	Leu	Gln	Asp	
					355				360					365			
30	ttc	tgt	ggc	tgt	gag	gtg	gct	ctg	gac	ctg	tgg	gaa	gac	ttc	agc	ctc	1263
	Phe	Cys	Gly	Cys	Glu	Val	Ala	Leu	Asp	Leu	Trp	Glu	Asp	Phe	Ser	Leu	
			370					375						380			
35	tgt	aga	gaa	ggg	cag	aga	gaa	tgg	gtc	atc	cag	aag	atc	cac	gag	tcc	1311
	Cys	Arg	Glu	Gly	Gln	Arg	Glu	Trp	Val	Ile	Gln	Lys	Ile	His	Glu	Ser	
		385					390					395					
40	cag	ttc	atc	att	gtg	gtt	tgt	tcc	aaa	ggg	atg	aag	tac	ttt	gtg	gac	1359
	Gln	Phe	Ile	Ile	Val	Val	Cys	Ser	Lys	Gly	Met	Lys	Tyr	Phe	Val	Asp	
		400				405					410						
	aag	aag	aac	tac	aaa	cac	aaa	gga	ggg	ggc	cga	ggc	tcg	ggg	aaa	gga	1407
	Lys	Lys	Asn	Tyr	Lys	His	Lys	Gly	Gly	Gly	Arg	Gly	Ser	Gly	Lys	Gly	
	415					420				425					430		
45	gag	ctc	ttc	ctg	gtg	gcg	gtg	tca	gcc	att	gcc	gaa	aag	ctc	cgc	cag	1455
	Glu	Leu	Phe	Leu	Val	Ala	Val	Ser	Ala	Ile	Ala	Glu	Lys	Leu	Arg	Gln	
				435					440						445		
50	gcc	aag	cag	agt	tcg	tcc	gcg	gcg	ctc	agc	aag	ttt	atc	gcc	gtc	tac	1503
	Ala	Lys	Gln	Ser	Ser	Ser	Ala	Ala	Leu	Ser	Lys	Phe	Ile	Ala	Val	Tyr	
			450						455				460				
	ttt	gat	tat	tcc	tcg	gag	gga	gac	gtc	ccc	ggg	atc	cta	gac	ctg	agt	1551
	Phe	Asp	Tyr	Ser	Cys	Glu	Gly	Asp	Val	Pro	Gly	Ile	Leu	Asp	Leu	Ser	
		465					470					475					
55	acc	aag	tac	aga	ctc	atg	gac	aat	ctt	cct	cag	ctc	tgt	tcc	cac	ctg	1599
	Thr	Lys	Tyr	Arg	Leu	Met	Asp	Asn	Leu	Pro	Gln	Leu	Cys	Ser	His	Leu	
		480					485				490						

	cac tcc cga gac cac ggc ctc cag gag ccg ggg cag cac acg cga cag	1647
	His Ser Arg Asp His Gly Leu Gln Glu Pro Gly Gln His Thr Arg Gln	
	495 500 505 510	
5	ggc agc aga agg aac tac ttc cgg agc aag tca ggc cgg tcc cta tac	1695
	Gly Ser Arg Arg Asn Tyr Phe Arg Ser Lys Ser Gly Arg Ser Leu Tyr	
	515 520 525	
10	gtc gcc att tgc aac atg cac cag ttt att gac gag gag ccc gac tgg	1743
	Val Ala Ile Cys Asn Met His Gln Phe Ile Asp Glu Glu Pro Asp Trp	
	530 535 540	
15	ttc gaa aag cag ttc gtt ccc ttc cat cct cct cca ctg cgc tac cgg	1791
	Phe Glu Lys Gln Phe Val Pro Phe His Pro Pro Pro Leu Arg Tyr Arg	
	545 550 555	
20	gag cca gtc ttg gag aaa ttt gat tcg ggc ttg gtt tta aat gat gtc	1839
	Glu Pro Val Leu Glu Lys Phe Asp Ser Gly Leu Val Leu Asn Asp Val	
	560 565 570	
25	atg tgc aaa cca ggg cct gag agt gac ttc tgc cta aag gta gag gcg	1887
	Met Cys Lys Pro Gly Pro Glu Ser Asp Phe Cys Leu Lys Val Glu Ala	
	575 580 585 590	
30	gct gtt ctt ggg gca acc gga cca gcc gac tcc cag cac gag agt cag	1935
	Ala Val Leu Gly Ala Thr Gly Pro Ala Asp Ser Gln His Glu Ser Gln	
	595 600 605	
35	cat ggg ggc ctg gac caa gac ggg gag gcc cgg cct gcc ctt gac ggt	1983
	His Gly Gly Leu Asp Gln Asp Gly Glu Ala Arg Pro Ala Leu Asp Gly	
	610 615 620	
40	agc gcc gcc ctg caa ccc ctg ctg cac acg gtg aaa gcc ggc agc ccc	2031
	Ser Ala Ala Leu Gln Pro Leu Leu His Thr Val Lys Ala Gly Ser Pro	
	625 630 635	
45	tcg gac atg ccg cgg gac tca ggc atc tat gac tcg tct gtg ccc tca	2079
	Ser Asp Met Pro Arg Asp Ser Gly Ile Tyr Asp Ser Ser Val Pro Ser	
	640 645 650	
50	tcc gag ctg tct ctg cca ctg atg gaa gga ctc tcg acg gac cag aca	2127
	Ser Glu Leu Ser Leu Pro Leu Met Glu Gly Leu Ser Thr Asp Gln Thr	
	655 660 665 670	
55	gaa acg tct tcc ctg acg gag agc gtg tcc tcc tct tca ggc ctg ggt	2175
	Glu Thr Ser Ser Leu Thr Glu Ser Val Ser Ser Ser Ser Gly Leu Gly	
	675 680 685	
60	gag gag gaa cct cct gcc ctt cct tcc aag ctc ctc tct tct ggg tca	2223
	Glu Glu Glu Pro Pro Ala Leu Pro Ser Lys Leu Leu Ser Ser Gly Ser	
	690 695 700	
65	tgc aaa gca gat ctt ggt tgc cgc agc tac act gat gaa ctc cac gcg	2271
	Cys Lys Ala Asp Leu Gly Cys Arg Ser Tyr Thr Asp Glu Leu His Ala	
	705 710 715	
70	gtc gcc cct ttg taacaaaacg aaagagtcta agcattgccca ctttagctgc	2323
	Val Ala Pro Leu	
	720	

tgcctccctc tgattcccca gctcatctcc ctggttgcat ggcccacttg gagctgaggt 2383  
 5 ctcatacaag gatatttggga gtgaaatgct ggccagtact tgttctccct tgccccaacc 2443  
 ctttaccgga tatcttgaca aactctccaa ttttctaaaa tgatatggag ctctgaaagg 2503  
 catgtccata aggtctgaca acagcttgcc aaatttggtt agtccttgga tcagagcctg 2563  
 10 ttgtgggagg tagggaggaa atatgtaaaag aaaaacagga agatacctgc actaatcatt 2623  
 cagacttcat tgagctctgc aaactttgcc tgtttgctat tggctacctt gatttgaaat 2683  
 gctttgtgaa aaaaggcact tttaacatca tagccacaga aatcaagtgc cagtctatct 2743  
 15 ggaatccatg ttgtattgca gataatgttc tcatttattt ttg 2786  
  
 MAPWLQLCSVFFTVNACLNGSQLAVAAGGSGRAXGADTCSWXGVGPASRNSGLYNITPKYDNCTTYLNPVGK  
 20 HVIADAQNITISQYACHDQVAVTILWSPGALGIEFLKGFRVILEELKSEGRQXQQLILKDPKQXNSSFRTG  
 MESQPxLNMKFETDYFVRLSFSFIKNESNYHPFFFRTRACDLLLPDNLACKPFWKPRNLNISQHGSDMQVS  
 FDHAPHNFGFRFFYLHYKLKHEGPFKRKTCKQEQTTEMTSCLLQNVSPGDYIIELVDDTNTTRKVMHYALKP  
 VHSPWAGPIRAVAITVPLVVISAFATLFTVMCRKKQENIYSHLDEESSESSTYTAALPRERLRPRPKVFLC  
 YSSKDGQNMNVVQCFAFLQDFCGCEVALDLWEDFSLCREGQREWVIQKIHSQFIIIVVCSKGMKYFVDK  
 25 NYKHKGGRGSGKGEFLVAVSAIAEKLQAKQSSSAALSKFIAVYFDYSCGDVPGILDSTKYRLMDNLP  
 QLCSHLSRDHGLQEPGQHTROGSRNRYFRSKSGRSLYVAICNMHQFIDEEDWFEKQVFPFHPPPLRYREP  
 VLEKFDGLVNDVMCKPGPESDFCLKVEAAVLGATGPADSQHESQHGGLDQDGEARPALDGSAAALQPLHT  
 VKAGSPDMPRDSGIYDSSVPSSELSLPLMEGLSTDQTETSSLTESVSSSSGLGEEPPALPSKLLSSGCK  
 ADLGCRSYTDELHAVAPL.  
  
 30 Reverse translation of primate, e.g., human, DCRS8 (SEQ ID NO: 15):  
 atggcnccnt ggytncaryt ntgywsngtn ttyttyacng tnaaygcntg yytnaayggn 60  
 35 wsnrcarytng cngtngcngc ngngngnwns ggnmgngcnn nngngngnga yacntgywsn 120  
 tggnnngngng tngngcngc nwsnmgnaay wsngngnynt ayaayathac nttyaartay 180  
 gayaaytgga cnacntayyt naayccngtn ggnaarcayg tnathgnga ygcncaraay 240  
 40 athacnathw sncartaygc ntgycaygay cargtngcng tnacnathyt ntggwsnccn 300  
 ggngcnytn gnatggartt yytnaarggn ttmngngtna thytngarga rytnaarwsn 360  
 45 gargngmgnc arnnncarca rytnathytn aargayccna arcarnnnaa ywsnwsntty 420  
 aarmgnacng gnatggarws ncarccnnnn ytnaayatga arttygarac ngaytaytty 480  
 gtnmgnytnw snttywsntt yathaaraay garwsnaayt aycayccntt yttyttymgn 540  
 50 acnmngngent gygayytnyt nytnrcarccn gayaayytng cntgyaarcc nttytggaar 600  
 ccnmgnaayy tnaayathws ncarcayggn wsngayatgc argtnwsntt ygaycaygn 660  
 55 ccncayaayt tyggnttymg nttyttytay ytncaytaya arytnarca ygargngccn 720  
 ttyaarmgna aracntgyaa rcargarcar acnacngara tgacnwsntg yytnytnar 780  
 aaygtnwsnc cngngayta yathathgar ytngtngayg ayacnaayac nacnmgnaar 840



gtnatgcayt aygcnytnaa rccngtncay wsnccntggg cnggnccnat hmgngcngtn 900  
 5 gcnathacng tncnnytngt ngtnathwsn gcnttygcna cnytnntyac ngtnatgtgy 960  
 mgnaaraarc arcargaraa yathtaywsn cayytngayg argarwsnws ngarwsnwsn 1020  
 acntayacng cngcnytncc nmngngarmgn ytnmgncnm gncnaargt nttyytnatgy 1080  
 10 taywsnwsna argayggna raaycaytg aaygtngtnc artgyttygc ntayttyytn 1140  
 cargayttyt gyggntgyga rgtngcnytn gayytnntggg argayttyw nytnngymgn 1200  
 garggncarm gngartgggt nathcaraar athcaygarw sncarttyat hathgtngtn 1260  
 15 tgywsnaarg gnatgaarta yttygtngay aaraaraayt ayaarcayaa rggngggngn 1320  
 mgnggnwsng gnaarggnga rytnttyytn gtngcngtnw sngcnathgc ngaraarytn 1380  
 20 mgncargcna arcarwsnws nwngcngcn ytnwsnaart tyathgcngt ntayttygay 1440  
 taywsntgyg arggngaygt nccnggnath ytngayytnw snacnaarta ymgnytnatg 1500  
 gayaayytn cncarytntg ywsncayytn caywsnmngn aycayggnyt ncargarccn 1560  
 25 ggncarcaya cnmgncargg nwsnmgnmgn aaytaytym gnwsnaarws nggnmgwnsn 1620  
 ytntaygtng cnathgyaa yatgcaycar ttyathgayg argarccnga ytggttygar 1680  
 30 aarcarttyg tncnttyca yccnccnccn ytnmgntaym gngarccngt nytnagaraar 1740  
 ttygaywsng gnytngtnyt naaygaygt atgtgyaarc cnggnccnga rwsngaytty 1800  
 35 tgyytnaarg tngargcngc ngtnytnngn gcnacnggnc cngcngayws ncarcaygar 1860  
 wsnarcayg gnggnytna ycargaygg gargcnmgnc cngcnytna yggwnsngcn 1920  
 gcnytncaarc cnytnytnca yacngtnaar gcnggnwsnc cnwsngayat gccnmngay 1980  
 40 wsggnatht aygaywsnws ngtnccnwsn wsgarytnw snytnccnyt natggarggn 2040  
 ytnwsnacng aycaracnga racnwsnwsn ytnacngarw sngtnwsnws nwsnwsngn 2100  
 ytnggngarg argarccncc ngcnytnccn wsnaarytny tnwsnwsngg nwsntgyaar 2160  
 45 gcngayytn gntgymgnws ntayacngay garytncayg cngtngcncc nytn 2214

50 Table 4: Nucleotide and polypeptide sequences of DNAX Cytokine Receptor Subunit like  
 embodiments (DCRS9). Primate, e.g., human, embodiment (see SEQ ID NO: 16 and 17).  
 Predicted signal sequence indicated, but may vary by a few positions and depending upon cell  
 type.

55 atg ggg agc tcc aga ctg gca gcc ctg ctc ctg cct ctc ctc ctc ata 48  
 Met Gly Ser Ser Arg Leu Ala Ala Leu Leu Leu Pro Leu Leu Leu Ile  
 -20 -15 -10

27.

	gtc atc gac ctc tct gac tct gct ggg att ggc ttt cgc cac ctg ccc	96
	Val Ile Asp Leu Ser Asp Ser Ala Gly Ile Gly Phe Arg His Leu Pro	
	-5 -1 1 5	
5	cac tgg aac acc cgc tgt cct ctg gcc tcc cac acg gaa gtt ctg cct	144
	His Trp Asn Thr Arg Cys Pro Leu Ala Ser His Thr Glu Val Leu Pro	
	10 15 20 25	
10	ata tcc ctt gcc gca cct ggt ggg ccc tct tct cca caa agc ctt ggt	192
	Ile Ser Leu Ala Ala Pro Gly Gly Pro Ser Ser Pro Gln Ser Leu Gly	
	30 35 40	
15	gtg tgc gag tct ggc act gtt ccc gct gtt tgt gcc agc atc tgc tgt	240
	Val Cys Glu Ser Gly Thr Val Pro Ala Val Cys Ala Ser Ile Cys Cys	
	45 50 55	
20	cag gtg gct cag gtc ttc aac ggg gcc tct tcc acc tcc tgg tgc aga	288
	Gln Val Ala Gln Val Phe Asn Gly Ala Ser Ser Thr Ser Trp Cys Arg	
	60 65 70	
25	aat cca aaa agt ctt cca cat tca agt tct ata gga gac aca aga tgc	336
	Asn Pro Lys Ser Leu Pro His Ser Ser Ser Ile Gly Asp Thr Arg Cys	
	75 80 85	
30	cag cac ctg ctc aga gga agc tgc tgc ctc gtc gtc acc tgt ctg aga	384
	Gln His Leu Leu Arg Gly Ser Cys Cys Leu Val Val Thr Cys Leu Arg	
	90 95 100 105	
35	aga gcc atc aca ttt cca tcc cct ccc cag aca tct ccc aca agg gac	432
	Arg Ala Ile Thr Phe Pro Ser Pro Pro Gln Thr Ser Pro Thr Arg Asp	
	110 115 120	
40	ttc gct cta aaa gga ccc aac ctt cgg atc cag aga cat ggg aaa gtc	480
	Phe Ala Leu Lys Gly Pro Asn Leu Arg Ile Gln Arg His Gly Lys Val	
	125 130 135	
45	ttc cca gat tgg act cac aaa ggc atg gag gtg ggc act ggg tac aac	528
	Phe Pro Asp Trp Thr His Lys Gly Met Glu Val Gly Thr Gly Tyr Asn	
	140 145 150	
50	agg aga tgg gtt cag ctg agt ggt gga ccc gag ttc tcc ttt gat ttg	576
	Arg Arg Trp Val Gln Leu Ser Gly Gly Pro Glu Phe Ser Phe Asp Leu	
	155 160 165	
55	ctg cct gag gcc cgg gct att cgg gtg acc ata tct tca ggc cct gag	624
	Leu Pro Glu Ala Arg Ala Ile Arg Val Thr Ile Ser Ser Gly Pro Glu	
	170 175 180 185	
60	gtc agc gtg cgt ctt tgt cac cag tgg gca ctg gag tgt gaa gag ctg	672
	Val Ser Val Arg Leu Cys His Gln Trp Ala Leu Glu Cys Glu Glu Leu	
	190 195 200	
65	agc agt ccc tat gat gtc cag aaa att gtg tct ggg ggc cac act gta	720
	Ser Ser Pro Tyr Asp Val Gln Lys Ile Val Ser Gly Gly His Thr Val	
	205 210 215	
70	gag ctg cct tat gaa ttc ctt ctg ccc tgt ctg tgc ata gag gca tcc	768
	Glu Leu Pro Tyr Glu Phe Leu Leu Pro Cys Leu Cys Ile Glu Ala Ser	
	220 225 230	

5	tac	ctg	caa	gag	gac	act	gtg	agg	cgc	aaa	aaa	tgt	ccc	ttc	cag	agc	816
	Tyr	Leu	Gln	Glu	Asp	Thr	Val	Arg	Arg	Lys	Lys	Cys	Pro	Phe	Gln	Ser	
	235						240					245					
10	tgg	cca	gaa	gcc	tat	ggc	tcg	gac	ttc	tgg	aag	tca	gtg	cac	ttc	act	864
	Trp	Pro	Glu	Ala	Tyr	Gly	Ser	Asp	Phe	Trp	Lys	Ser	Val	His	Phe	Thr	
	250					255					260					265	
15	gac	tac	agc	cag	cac	act	cag	atg	gtc	atg	gcc	ctg	aca	ctc	cgc	tgc	912
	Asp	Tyr	Ser	Gln	His	Thr	Gln	Met	Val	Met	Ala	Leu	Thr	Leu	Arg	Cys	
					270					275					280		
20	cca	ctg	aag	ctg	gaa	gct	gcc	ctc	tgc	cag	agg	cac	gac	tgg	cat	acc	960
	Pro	Leu	Lys	Leu	Glu	Ala	Ala	Leu	Cys	Gln	Arg	His	Asp	Trp	His	Thr	
				285					290					295			
25	ctt	tgc	aaa	gac	ctc	ccg	aat	gcc	acg	gct	cga	gag	tca	gat	ggg	tgg	1008
	Leu	Cys	Lys	Asp	Leu	Pro	Asn	Ala	Thr	Ala	Arg	Glu	Ser	Asp	Gly	Trp	
			300					305					310				
30	tat	gtt	ttg	gag	aag	gtg	gac	ctg	cac	ccc	cag	ctc	tgc	ttc	aag	gta	1056
	Tyr	Val	Leu	Glu	Lys	Val	Asp	Leu	His	Pro	Gln	Leu	Cys	Phe	Lys	Val	
	315						320						325				
35	caa	cca	tgg	ttc	tct	ttt	gga	aac	agc	agc	cat	gtt	gaa	tgc	ccc	cac	1104
	Gln	Pro	Trp	Phe	Ser	Phe	Gly	Asn	Ser	Ser	His	Val	Glu	Cys	Pro	His	
	330					335					340				345		
40	cag	act	ggg	tct	ctc	aca	tcc	tgg	aat	gta	agc	atg	gat	acc	caa	gcc	1152
	Gln	Thr	Gly	Ser	Leu	Thr	Ser	Trp	Asn	Val	Ser	Met	Asp	Thr	Gln	Ala	
					350					355					360		
45	cag	cag	ctg	att	ctt	cac	ttc	tcc	tca	aga	atg	cat	gcc	acc	ttc	agt	1200
	Gln	Gln	Leu	Ile	Leu	His	Phe	Ser	Ser	Arg	Met	His	Ala	Thr	Phe	Ser	
				365					370						375		
50	gct	gcc	tgg	agc	ctc	cca	ggc	ttg	ggg	cag	gac	act	ttg	gtg	ccc	ccc	1248
	Ala	Ala	Trp	Ser	Leu	Pro	Gly	Leu	Gly	Gln	Asp	Thr	Leu	Val	Pro	Pro	
			380					385						390			
55	gtg	tac	act	gtc	agc	cag	gtg	tgg	cgg	tca	gat	gtc	cag	ttt	gcc	tgg	1296
	Val	Tyr	Thr	Val	Ser	Gln	Val	Trp	Arg	Ser	Asp	Val	Gln	Phe	Ala	Trp	
	395						400					405					
60	aag	cac	ctc	ttg	tgt	cca	gat	gtc	tct	tac	aga	cac	ctg	ggg	ctc	ttg	1344
	Lys	His	Leu	Leu	Cys	Pro	Asp	Val	Ser	Tyr	Arg	His	Leu	Gly	Leu	Leu	
	410					415					420				425		
65	atc	ctg	gca	ctg	ctg	gcc	ctc	ctc	acc	cta	ctg	ggg	gtt	gtt	ctg	gcc	1392
	Ile	Leu	Ala	Leu	Leu	Ala	Leu	Leu	Thr	Leu	Leu	Gly	Val	Val	Leu	Ala	
					430					435					440		
70	ctc	acc	tgc	cgg	cgc	cca	cag	tca	ggc	ccg	ggc	cca	gcg	cgg	cca	gtg	1440
	Leu	Thr	Cys	Arg	Arg	Pro	Gln	Ser	Gly	Pro	Gly	Pro	Ala	Arg	Pro	Val	
				445					450					455			

	ctc ctc ctg cac gcg gcg gac tcg gag gcg cag cgg cgc ctg gtg gga	1488
	Leu Leu Leu His Ala Ala Asp Ser Glu Ala Gln Arg Arg Leu Val Gly	
	460 465 470	
5	gcg ctg gct gaa ctg cta cgg gca gcg ctg ggc ggc ggg cgc gac gtg	1536
	Ala Leu Ala Glu Leu Leu Arg Ala Ala Leu Gly Gly Gly Arg Asp Val	
	475 480 485	
10	atc gtg gac ctg tgg gag ggg agg cac gtg gcg cgc gtg ggc ccg ctg	1584
	Ile Val Asp Leu Trp Glu Gly Arg His Val Ala Arg Val Gly Pro Leu	
	490 495 500 505	
15	ccg tgg ctc tgg gcg gcg cgg acg cgc gta gcg cgg gag cag ggc act	1632
	Pro Trp Leu Trp Ala Ala Arg Thr Arg Val Ala Arg Glu Gln Gly Thr	
	510 515 520	
20	gtg ctg ctg ctg tgg agc ggc gcc gac ctt cgc ccg gtc agc ggc ccc	1680
	Val Leu Leu Leu Trp Ser Gly Ala Asp Leu Arg Pro Val Ser Gly Pro	
	525 530 535	
	gac ccc cgc gcc gcg ccc ctg ctc gcc ctg ctc cac gct gcc ccg cgc	1728
	Asp Pro Arg Ala Ala Pro Leu Leu Ala Leu Leu His Ala Ala Pro Arg	
	540 545 550	
25	ccg ctg ctg ctg ctc gct tac ttc agt cgc ctc tgc gcc aag ggc gac	1776
	Pro Leu Leu Leu Leu Ala Tyr Phe Ser Arg Leu Cys Ala Lys Gly Asp	
	555 560 565	
30	atc ccc ccg ccg ctg cgc gcc ctg ccg cgc tac cgc ctg ctg cgc gac	1824
	Ile Pro Pro Pro Leu Arg Ala Leu Pro Arg Tyr Arg Leu Leu Arg Asp	
	570 575 580 585	
35	ctg ccg cgt ctg ctg cgg gcg ctg gac gcg cgg cct ttc gca gag gcc	1872
	Leu Pro Arg Leu Leu Arg Ala Leu Asp Ala Arg Pro Phe Ala Glu Ala	
	590 595 600	
40	acc agc tgg ggc cgc ctt ggg gcg cgg cag cgc agg cag agc cgc cta	1920
	Thr Ser Trp Gly Arg Leu Gly Ala Arg Gln Arg Arg Gln Ser Arg Leu	
	605 610 615	
	gag ctg tgc agc cgg ctc gaa cga gag gcc gcc cga ctt gca gac cta	1968
	Glu Leu Cys Ser Arg Leu Glu Arg Glu Ala Ala Arg Leu Ala Asp Leu	
	620 625 630	
45	ggt tgagcagagc tccaccgcag tcccgggtgt ctgcggccgc t	2012
	Gly	
50	MGSSRLAALLPLLLIVIDLSDSAGIGFRHLPHWNTRCPLASHTEVLPISLAAPGGPSSPQSLGVCESGTVP	
	AVCASICQVAQVFNGASSTSWCRNPKSLPHSSSIGDTRCQHLLRGSCCLVVTCLRRAITFPSPPTSPTRD	
	FALKGPNLRIQRHGKVPDPWTHKGMEVGTYGNNRRWVQLSGGPEFSFDLLPEARAIRVTISSGPEVSVRLCHQ	
	WALECEBELSSPYDVQKIVSGGHTVELPYEFLLPCLCIEASYLQEDTVRRKKCPFSWPEAYGSDFWKSVHFT	
	DYSQHTQMVMALTLRCPLKLEAALCQRHDWHTLCKDLPNATARESDGWYVLEKVDLHPQLCFKVQPWFSFGN	
	SSHVECPHQTSLSWNVSMDTQAQQLILHFSSRMHATFSAAWSLPGLGQDTLVPPVYTVSVQVWRSVDQFAW	
	KHLCPDVSYRHLGLLILALLALLTLGVLALTCCRPGSGPGPARPVLLLHAADSEAQRRLVGALAEELLRA	
55	ALGGGRDVIVDLWEGRHVARVGPLPWLWAARTRVAREQGTVLLWLGADLRPVSGPDPRAPLLALLHAAPR	
	PLLLLAYFSRLCAKGDIPPLRALPRYRLRLDLPRLLRALDARPF AEATSWGRLGARQRRQSRLELCSRLER	
	EAARLADLG.	

Reverse translation of primate, e.g., human, DCRS9 (SEQ ID NO: 18):

5 atgggnwsnw snmgnytn gc ngenytnytn ytnccnytny tnytnathgt nathgayytn 60  
 wsn gaywsng cnggnathgg nttymgncay ytnccncayt ggaayacnmg ntgyccnytn 120  
 gc nwsncaya cngargtnyt nccnathwsn ytn gcngcnc cnggnggncc nwsnwsnccn 180  
 10 carwsnytn gngtntgyga rwsnggnacn gtncngcng tntgygc nws nathtgytgy 240  
 cargtn gcnc argtnttyaa yggngcnwsn wsnacnwsnt ggtgymgnaa yccnaarwsn 300  
 15 ytnccncayw snwsnwsnat hggngayacn mgntgy carc ayytnytnmg nggnwsntgy 360  
 tgyytn gtn gtnacntgyt nmgnmgngcn athacnttyc cnwsncncnc ncaracnwsn 420  
 ccnacnmng aytygcn yt naarggnccn ayytnmgna thcarmgna yggnaargtn 480  
 20 ttyccngayt ggacncayaa rggnatggar gtnggnacng gntayaaymg nmngtgggt 540  
 carytnwsng gnggnccnga rtywsntty gayytnytn cngargcnmg ngcnathmgn 600  
 gtnacnathw snwsnggncc ngargtnwsn gtnmgnytn gycaycartg ggcnytn gar 660  
 25 tgygargary tnwsnwsncc ntaygaygn caraarathg tnwsngngng ncayacngtn 720  
 garytnccnt aygarttyt nytnccntgy ytntgyathg argcnwsnta yytnccarg 780  
 30 gayacngtnm gnmgnaaraa rtgyccntty carwsntggc cngargenta ygg nwsngay 840  
 ttytggaarw sngtn caytt yacngaytay wsnarcaya cncaratggt natggcn ytn 900  
 acnytnmgnt gyccnytnaa rytngargcn gcnytn gyc armgncayga ytggcayacn 960  
 35 ytntgyaarg ayytnccnaa ygcncngcn mgngarwsng ayggntggta ygtnytn gar 1020  
 aargtn gayy tncayccna rytntgytty aargtn carc cntggtyws nttyggnaay 1080  
 40 wsnwsncayg tngartgycc ncaycaracn ggnwsnytna cnwsntggaa ygtnwsnatg 1140  
 gayacncarg cncarcaryt nathytn cay ttwsnwsnm gnatgcaygc nacnttywsn 1200  
 gcngcntggw snytnccngg nytnngncar gayacnytn tncncncgt ntayacngtn 1260  
 45 wsnargtn ggmgnwsnga ygtncartty gcntggaarc ayytnytn g yccngaygn 1320  
 wsntaymgnc ayytnngny nytnathytn gcnytnytn cnytnytnac nytnytnngn 1380  
 50 gtngtnytn gnytnacntg ymgngmgnccn carwsnggnc cnggnccngc nmgnccngtn 1440  
 ytnytnytn aygcngcnga ywsngargcn carmgmgn ytn gngngc nytn gncgar 1500  
 ytnytnmgng cngcn ytn ngngngmgn gaygtnathg tngayytn ggar gngmgn 1560  
 55 caygtn gnm gngtnngncc nytnccntgg ytntgggcng cnmgnacnmg ngtn gcnmgn 1620  
 garcarggna cngtnytn nytn ggnwsn gngcngayy tnmgnccngt nwsnggnccn 1680

gayccnmngng cngcncnyt nytngcnytn ytncaygeng cncnmgnc nytnytnytn 1740  
 ytngcntayt tywsnmgnyt ntgygcnaar ggngayathc cncncncnyt nmngcnytn 1800  
 5 ccnmgntaym gnytnytnmg ngayytnccn mgnytnytnm gngcnytna ygcnmgnccn 1860  
 ttygcngarg cnacnwsntg gggnmgnytn ggngcnmgnc armgnmgna rwsnmgnytn 1920  
 garytntgyw snmgnytna rmngngargcn gcnmgnytn cngayytnng n 1971  
 10

Rodent, e.g., mouse, embodiment (see SEQ ID NO: 19 and 20). Predicted signal sequence indicated, but may vary by a few positions and depending upon cell type.

15 cagctccggg ccaggccctg ctgccctctt gcagacagga aagacatggt ctctgcgccc 60  
 tgatcctaca gaagctc atg ggg agc ccc aga ctg gca gcc ttg ctc ctg 110  
 Met Gly Ser Pro Arg Leu Ala Ala Leu Leu Leu  
 -20 -15  
 20 tct ctc ccg cta ctg ctc atc ggc ctc gct gtg tct gct cgg gtt gcc 158  
 Ser Leu Pro Leu Leu Leu Ile Gly Leu Ala Val Ser Ala Arg Val Ala  
 -10 -5 -1 1  
 25 tgc ccc tgc ctg cgg agt tgg acc agc cac tgt ctc ctg gcc tac cgt 206  
 Cys Pro Cys Leu Arg Ser Trp Thr Ser His Cys Leu Leu Ala Tyr Arg  
 5 10 15 20  
 30 gtg gat aaa cgt ttt gct ggc ctt cag tgg ggc tgg ttc cct ctc ttg 254  
 Val Asp Lys Arg Phe Ala Gly Leu Gln Trp Gly Trp Phe Pro Leu Leu  
 25 30 35  
 35 gtg agg aaa tct aaa agt cct cct aaa ttt gaa gac tat tgg agg cac 302  
 Val Arg Lys Ser Lys Ser Pro Pro Lys Phe Glu Asp Tyr Trp Arg His  
 40 45 50  
 40 agg aca cca gca tcc ttc cag agg aag ctg cta ggc agc cct tcc ctg 350  
 Arg Thr Pro Ala Ser Phe Gln Arg Lys Leu Leu Gly Ser Pro Ser Leu  
 55 60 65  
 45 tct gag gaa agc cat cga att tcc atc ccc tcc tca gcc atc tcc cac 398  
 Ser Glu Glu Ser His Arg Ile Ser Ile Pro Ser Ser Ala Ile Ser His  
 70 75 80  
 50 aga ggc caa cgc acc aaa agg gcc cag cct tca gct gca gaa gga aga 446  
 Arg Gly Gln Arg Thr Lys Arg Ala Gln Pro Ser Ala Ala Glu Gly Arg  
 85 90 95 100  
 55 gaa cat ctc cct gaa gca ggg tca caa aag tgt gga gga cct gaa ttc 494  
 Glu His Leu Pro Glu Ala Gly Ser Gln Lys Cys Gly Gly Pro Glu Phe  
 105 110 115  
 tcc ttt gat ttg ctg ccc gag gtg cag gct gtt cgg gtg act att cct 542  
 Ser Phe Asp Leu Leu Pro Glu Val Gln Ala Val Arg Val Thr Ile Pro  
 120 125 130

	gca ggc ccc aag gca cgt gtg cgc ctt tgt tat cag tgg gca ctg gaa	590
	Ala Gly Pro Lys Ala Arg Val Arg Leu Cys Tyr Gln Trp Ala Leu Glu	
	135 140 145	
5	tgt gaa gac ttg agt agc cct ttt gat acc cag aaa att gtg tct gga	638
	Cys Glu Asp Leu Ser Ser Pro Phe Asp Thr Gln Lys Ile Val Ser Gly	
	150 155 160	
10	ggg cac act gta gac ctg cct tat gaa ttc ctt ctg ccc tgc atg tgc	686
	Gly His Thr Val Asp Leu Pro Tyr Glu Phe Leu Leu Pro Cys Met Cys	
	165 170 175 180	
15	ata gag gcc tcc tac ctg caa gag gac act gtg agg cgc aaa agt gtc	734
	Ile Glu Ala Ser Tyr Leu Gln Glu Asp Thr Val Arg Arg Lys Ser Val	
	185 190 195	
20	cct tcc aga gct ggc ctg aag ctt atg gct cag act tct ggc agt caa	782
	Pro Ser Arg Ala Gly Leu Lys Leu Met Ala Gln Thr Ser Gly Ser Gln	
	200 205 210	
	tac gct tca ctg act aca gcc agc ac	808
	Tyr Ala Ser Leu Thr Thr Ala Ser	
	215 220	
25	MGSPRLAALLLSLPLLLIGLAVSARVACPCLRSWTSCHLLAYRVDKRFAGLQGWGFPLLVRKSKSPPKFEDY WRHRTPASFQRKLLGSPSLSEBSHRISIPSSAISHRGQRTKRAQPSAAEGREHLPEAGSQKCGGPEFSFDLL PEVQAVRVTIPAGPKARVRLCYQWALECEDLSSPFDQKIVSGGHTVDLPYEFLLPCMCIEASYLQEDTVRR KSVPSRAGLKLMAQTSQSQYASLTTAS	
30	Reverse translation of rodent, e.g., mouse, DCRS9 (SEQ ID NO: 21):	
	atgggnwsnc cnmgnytngc ngcnytnytn ytnwsnytn cnytnytnyt nathggnytn	60
35	gcngtnwsng cnmgngtngc ntgyccntgy ytnmgwsnt ggacnwsnca ytgyytnytn	120
	gcntaymgng tngayaarmg nttygcnggn ytncartggg gntgggttycc nytnytnytn	180
40	mgnaarwsna arwsncchcc naarttygar gaytaytggm gncaymgna nccngcnwsn	240
	ttycarmgna arytnytnng nwsnccnwsn ytnwsngarg arwsncaymg nathwsnath	300
	ccnwsnwsng cnathwsnca ymgnggncar mgnaacnaarm gngcncarcc nwsngcngcn	360
45	garggnmgng arcayytnc ngargcnggn wscaraart gyggnggngcc ngarttywsn	420
	ttygayytny tncngargt ncargcngtn mgngtnacna thcngcngg nccnaargcn	480
50	mgngtnmgny tntgytayca rtgggcnytn gartgygarg ayytnwsnws nccnttygay	540
	acncaraara thgtnwsngg nggncayacn gtngayytnc cntaygartt yytnytnccn	600
	tgyatgtgya thgargcnws ntayytncar gargayacng tnmgnmgnaa rwsngtnccn	660
55	wsnmgngcng gnytnaaryt natggcncar acnwsnggnw sncartaygc nwsnytnacn	720
	acngcnwsn	729

Table 5: Nucleotide and polypeptide sequences of DNAX Cytokine Receptor Subunit like embodiments (DCRS10). Primate, e.g., human, embodiment (see SEQ ID NO: 22 and 23).

5	ttttgagcag aggccttccta ggctccgtag aaatttgcat acagcttcca cttcctgctt 60	
	cagagcctgt tcttctactt acctgggccc ggagaagggtg gagggagacg agaagccgcc 120	
10	gagagccgac taccctccgg gccagctctg tctgtccgtg gtggatctaa gaaactaga 179	
	atg aac cga agc att cct gtg gag gtt gat gaa tca gaa cca tac cca 227	
	Met Asn Arg Ser Ile Pro Val Glu Val Asp Glu Ser Glu Pro Tyr Pro	
	1 5 10 15	
15	agt cag ttg ctg aaa cca atc cca gaa tat tcc ccg gaa gag gaa tca 275	
	Ser Gln Leu Leu Lys Pro Ile Pro Glu Tyr Ser Pro Glu Glu Glu Ser	
	20 25 30	
20	gaa cca cct gct cca aat ata agg aac atg gca ccc aac agc ttg tct 323	
	Glu Pro Pro Ala Pro Asn Ile Arg Asn Met Ala Pro Asn Ser Leu Ser	
	35 40 45	
25	gca ccc aca atg ctt cac aat tcc tcc gga gac ttt tct caa gct cac 371	
	Ala Pro Thr Met Leu His Asn Ser Ser Gly Asp Phe Ser Gln Ala His	
	50 55 60	
30	tca acc ctg aaa ctt gca aat cac cag cgg cct gta tcc cgg cag gtc 419	
	Ser Thr Leu Lys Leu Ala Asn His Gln Arg Pro Val Ser Arg Gln Val	
	65 70 75 80	
35	acc tgc ctg cgc act caa gtt ctg gag gac agt gaa gac agt ttc tgc 467	
	Thr Cys Leu Arg Thr Gln Val Leu Glu Asp Ser Glu Asp Ser Phe Cys	
	85 90 95	
40	agg aga cac cca ggc ctg ggc aaa gct ttc cct tct ggg tgc tct gca 515	
	Arg Arg His Pro Gly Leu Gly Lys Ala Phe Pro Ser Gly Cys Ser Ala	
	100 105 110	
45	gtc agc gag cct gcg tct gag tct gtg gtt gga gcc ctc cct gca gag 563	
	Val Ser Glu Pro Ala Ser Glu Ser Val Val Gly Ala Leu Pro Ala Glu	
	115 120 125	
50	cat cag ttt tca ttt atg gaa aaa cgt aat caa tgg ctg gta tct cag 611	
	His Gln Phe Ser Phe Met Glu Lys Arg Asn Gln Trp Leu Val Ser Gln	
	130 135 140	
55	ctt tca gcg gct tct cct gac act ggc cat gac tca gac aaa tca gac 659	
	Leu Ser Ala Ala Ser Pro Asp Thr Gly His Asp Ser Asp Lys Ser Asp	
	145 150 155 160	
60	caa agt tta cct aat gcc tca gca gac tcc ttg ggc ggt agc cag gag 707	
	Gln Ser Leu Pro Asn Ala Ser Ala Asp Ser Leu Gly Gly Ser Gln Glu	
	165 170 175	
65	atg gtg caa cgg ccc cag cct cac agg aac cga gca ggc ctg gat ctg 755	
	Met Val Gln Arg Pro Gln Pro His Arg Asn Arg Ala Gly Leu Asp Leu	
	180 185 190	



	cca acc ata gac acg gga tat gat tcc cag ccc cag gat gtc ctg ggc	803
	Pro Thr Ile Asp Thr Gly Tyr Asp Ser Gln Pro Gln Asp Val Leu Gly	
	195 200 205	
5	atc agg cag ctg gaa agg ccc ctg ccc ctc acc tcc gtg tgt tac ccc	851
	Ile Arg Gln Leu Glu Arg Pro Leu Pro Leu Thr Ser Val Cys Tyr Pro	
	210 215 220	
10	cag gac ctc ccc aga cct ctc agg tcc agg gag ttc cct cag ttt gaa	899
	Gln Asp Leu Pro Arg Pro Leu Arg Ser Arg Glu Phe Pro Gln Phe Glu	
	225 230 235 240	
15	cct cag agg tat cca gca tgt gca cag atg ctg cct ccc aat ctt tcc	947
	Pro Gln Arg Tyr Pro Ala Cys Ala Gln Met Leu Pro Pro Asn Leu Ser	
	245 250 255	
20	cca cat gct cca tgg aac tat cat tac cat tgt cct gga agt ccc gat	995
	Pro His Ala Pro Trp Asn Tyr His Tyr His Cys Pro Gly Ser Pro Asp	
	260 265 270	
	cac cag gtg cca tat ggc cat gac tac cct cga gca gcc tac cag caa	1043
	His Gln Val Pro Tyr Gly His Asp Tyr Pro Arg Ala Ala Tyr Gln Gln	
	275 280 285	
25	gtg atc cag ccg gct ctg cct ggg cag ccc ctg cct gga gcc agt gtg	1091
	Val Ile Gln Pro Ala Leu Pro Gly Gln Pro Leu Pro Gly Ala Ser Val	
	290 295 300	
30	aga ggc ctg cac cct gtg cag aag gtt atc ctg aat tat ccc agc ccc	1139
	Arg Gly Leu His Pro Val Gln Lys Val Ile Leu Asn Tyr Pro Ser Pro	
	305 310 315 320	
35	tgg gac caa gaa gag agg ccc gca cag aga gac tgc tcc ttt ccg ggg	1187
	Trp Asp Gln Glu Arg Pro Ala Gln Arg Asp Cys Ser Phe Pro Gly	
	325 330 335	
40	ctt cca agg cac cag gac cag cca cat cac cag cca cct aat aga gct	1235
	Leu Pro Arg His Gln Asp Gln Pro His His Gln Pro Pro Asn Arg Ala	
	340 345 350	
	ggg gct cct ggg gag tcc ttg gag tgc cct gca gag ctg aga cca cag	1283
	Gly Ala Pro Gly Glu Ser Leu Glu Cys Pro Ala Glu Leu Arg Pro Gln	
	355 360 365	
45	gtt ccc cag cct ccg tcc cca gct gct gtg cct aga ccc cct agc aac	1331
	Val Pro Gln Pro Pro Ser Pro Ala Ala Val Pro Arg Pro Pro Ser Asn	
	370 375 380	
50	cct cca gcc aga gga act cta aaa aca agc aat ttg cca gaa gaa ttg	1379
	Pro Pro Ala Arg Gly Thr Leu Lys Thr Ser Asn Leu Pro Glu Glu Leu	
	385 390 395 400	
55	cgg aaa gtc ttt atc act tat tgc atg gac aca gct atg gag gtg gtg	1427
	Arg Lys Val Phe Ile Thr Tyr Ser Met Asp Thr Ala Met Glu Val Val	
	405 410 415	
	aaa ttc gtg aac ttt ttg ttg gta aat ggc ttc caa act gca att gac	1475
	Lys Phe Val Asn Phe Leu Leu Val Asn Gly Phe Gln Thr Ala Ile Asp	
	420 425 430	

ata ttt gag gat aga atc cga ggc att gat atc att aaa tgg atg gag 1523  
 Ile Phe Glu Asp Arg Ile Arg Gly Ile Asp Ile Ile Lys Trp Met Glu  
 435 440 445

5

cgc tac ctt agg gat aag acc gtg atg ata atc gta gca atc agc ccc 1571  
 Arg Tyr Leu Arg Asp Lys Thr Val Met Ile Ile Val Ala Ile Ser Pro  
 450 455 460

10

aaa tac aaa cag gac gtg gaa ggc gct gag tgc cag ctg gac gag gat 1619  
 Lys Tyr Lys Gln Asp Val Glu Gly Ala Glu Ser Gln Leu Asp Glu Asp  
 465 470 475 480

15

gag cat ggc tta cat act aag tac att cat cga atg atg cag att gag 1667  
 Glu His Gly Leu His Thr Lys Tyr Ile His Arg Met Met Gln Ile Glu  
 485 490 495

20

ttc ata aaa caa gga agc atg aat ttc aga ttc atc cct gtg ctc ttc 1715  
 Phe Ile Lys Gln Gly Ser Met Asn Phe Arg Phe Ile Pro Val Leu Phe  
 500 505 510

25

cca aat gct aag aag gag cat gtg ccc acc tgg ctt cag aac act cat 1763  
 Pro Asn Ala Lys Lys Glu His Val Pro Thr Trp Leu Gln Asn Thr His  
 515 520 525

gtc tac agc tgg ccc aag aat aaa aaa aac atc ctg ctg cgg ctg ctg 1811  
 Val Tyr Ser Trp Pro Lys Asn Lys Lys Asn Ile Leu Leu Arg Leu Leu  
 530 535 540

30

aga gag gaa gag tat gtg gct cct cca cgg ggg cct ctg ccc acc ctt 1859  
 Arg Glu Glu Glu Tyr Val Ala Pro Pro Arg Gly Pro Leu Pro Thr Leu  
 545 550 555 560

35

cag gtg gtt ccc ttg tgacaccgtt catccccaga tcaactgaggc caggccatgt 1914  
 Gln Val Val Pro Leu  
 565

40

ttggggcctt gttctgacag cattctggct gaggctggtc ggtagcactc ctggctgggtt 1974  
 tttttctgtt cctccccgag aggccctctg gccccagga aacctgttgt gcagagctct 2034  
 tccccggaga cctccacaca ccctggcttt gaagtggagt ctgtgactgc tctgcattct 2094

45

ctgcttttaa aaaaaccatt gcagggtcca gtgtcccata tgttcctcct gacagtttga 2154  
 tgtgtccatt ctgggcctct cagtgttag caagtagata atgtaaggga tgtggcagca 2214  
 aatggaaatg actacaaaca ctctcctatc aatcacttca ggctactttt atgagtttagc 2274

50

cagatgcttg tgtatcctca gaccaaactg attcatgtac aaataataaa atgtttactc 2334  
 ttttgtaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa 2377

5 MNR SIPVEVDESEEPYPSQLLKPIPEYSPEESEEPPAPNIRNMAPNSLSAPTMLHNSSGDFSQAHSTLKL ANH  
 QRPVSRQVTC LRTQVLEDSF CRRHPGLGKAPPSGCSAVSEPASES VVGALPAEHQFSFMEKRNQWLVSQ  
 LSAASPD TGHDSDKSDQSLPNASADSLGGSQEMVQR PQPHNRAGLDLPTIDTGYDSQPQDVLGIRQLERPL  
 10 PLTSVCYPQDLPRPLRSREFPQFEPQRY PACAQMLPPNLS PHAPWNYHYHCPGSPDHQVPYGHDPRAAYQQ  
 VIQPALPGQPLPGASVRGLHPVQKVILNYPSPWDQBERPAQRDCSFPGLPRHQDQPHHQP PNRAGAPGESLE  
 CPABLRPQVPQPPSPA AVPRPPSNPPARGTLKTSNLPEELRKVFITYSMDTAMEVVKFVNFLLVNGFQTAID  
 IFEDRIRGIDI I KW MERYLRDKTVMII VAISPKYKQDVEGAESQLDEDEHGLHTKYIHRMMQIEFIKQGS MN  
 FRFIPVLPNAPNAKEHVPTWLQNT HVYSWPKNKNILLRLLREEEYVAPPRGPLPTLQVVPL

Reverse translation of primate, e.g., human, DCRS10 (SEQ ID NO: 24):

15 atgaaymg nw snathccngt ngargtngay garwsngarc cntayccnws ncarytnytn 60  
 aarccnathc cngartayws nccngargar garwsngarc cncngcnc naayathmgn 120  
 aayatggcnc cnaaywsnyt nwsngcncn acnatgytnc ayaaywsnws ngngaytty 180  
 20 wsnrcargcnc aywsnacnyt naarytngcn aaycaycarm gncngtnws nmgnrcargtn 240  
 acntgyytnm gnacncargt nytngargay wsnrgargayw snttytgymg nmgnrcayccn 300  
 ggnytnngna argcnttycc nwsnggntgy wsnrgngtnw sngarccngc nwsngarwsn 360  
 25 gtngtnngng cnytnccngc ngarcaycar ttywsnttya tggaraarmg naaycartgg 420  
 ytngtnwsnc arytnwsngc ngcnwsncn gayacnggnc aygaywsnga yaarwsngay 480  
 30 carwsnytn cnaaygc nws ngcngaywsn ytngngngnw sncargarat ggtncarmgn 540  
 ccncarccnc aymgnaaymg ngcnggnytn gayytnccna cnathgayac nggntaygay 600  
 wsnrcarccnc argaygtnyt nggnathmgn carytngarm gncnytncc nytnacnwsn 660  
 35 gtntgytayc cncargayt nccnmgnccn ytnmgnwsnm gngarttycc ncarttygar 720  
 ccncarmgt aycngcngtg ygcncaratg ytnccncna ayytnwsncc ncaygcncn 780  
 40 tggaytayc aytaycaytg yccngnwsn ccngaycayc argtnccnta yggncaygay 840  
 tayccnmng cngcntayca rcargtnath carccngcny tncnggnc rccnytnccn 900  
 ggngcnwsng tnmnggnytn ncayccngtn caraargtna thytnaayta yccnwsncn 960  
 45 tggaycarg argarmgncc ngncarmgn gaytgywsnt tyccnggnytn nccnmgnay 1020  
 cargaycarc ncaycayca rccncnaay mgngcngng cncnggnga rwsnytngar 1080  
 50 tgyccngcng arytnmgncc ncargtnccn carccncnw sncngcngc ngtnccnmgn 1140  
 ccncnwsna aycncngc nmnggnacn ytnaaracnw snaaytncc ngargarytn 1200  
 mgnaargtn tyathacnta ywsnatggay acngcnatgg argtngtana rtyygtana 1260  
 55 tyytynytn tnaaygntt ycaracngcn athgayatht tygargaymg nathmngngn 1320  
 athgayatha thaartgat ggarmgtay ytnmngaya aracngtnat gathathgn 1380

genathwsnc cnaartayaa rcargaygtn gargngcng arwsncaryt ngaygargay 1440  
 garcayggny tncayacnaa rtayathcay mgnatgatgc arathgartt yathaarc 1500  
 5 ggnwsnatga ayttymgntt yathccngtn ytnttyccna aygcnaaraa rgarcaygtn 1560  
 ccnactggy tncaraayac ncaygntay wsntggccna araayaaraa raayathytn 1620  
 ytnmgnytny tnmngnarga rgartaygtn gcncncnmm gnggnccnyt nccnacnytn 1680  
 10 cargtngtnc cnytn 1695

Rodent, e.g., mouse, embodiment (see SEQ ID NO: 25 and 26).

15 cag gac ctc cct ggg cct ctg agg tcc agg gaa ttg cca cct cag ttt 48  
 Gln Asp Leu Pro Gly Pro Leu Arg Ser Arg Glu Leu Pro Pro Gln Phe  
 1 5 10 15  
 20 gaa ctt gag agg tat cca atg aac gcc cag ctg ctg ccg ccc cat cct 96  
 Glu Leu Glu Arg Tyr Pro Met Asn Ala Gln Leu Leu Pro Pro His Pro  
 20 25 30  
 25 tcc cca cag gcc cca tgg aac tgt cag tac tac tgc ccc gga ggg ccc 144  
 Ser Pro Gln Ala Pro Trp Asn Cys Gln Tyr Tyr Cys Pro Gly Gly Pro  
 35 40 45  
 30 tac cac cac cag gtg cca cac ggc cat ggc tac cct cca gca gca gcc 192  
 Tyr His His Gln Val Pro His Gly His Gly Tyr Pro Pro Ala Ala Ala  
 50 55 60  
 35 tac cag caa gta ctc cag cct gct ctg cct ggg cag gtc ctt cct ggg 240  
 Tyr Gln Gln Val Leu Gln Pro Ala Leu Pro Gly Gln Val Leu Pro Gly  
 65 70 75 80  
 gca agg gca aga ggc cca cgc cct gtg cag aag gtc atc ctg aat gac 288  
 Ala Arg Ala Arg Gly Pro Arg Pro Val Gln Lys Val Ile Leu Asn Asp  
 85 90 95  
 40 tcc agc ccc caa gac caa gaa gag aga cct gca cag aga gac ttc tct 336  
 Ser Ser Pro Gln Asp Gln Glu Glu Arg Pro Ala Gln Arg Asp Phe Ser  
 100 105 110  
 45 ttc ccg agg ctc ccg agg gac cag ctc tac cgc cca cca tct aat gga 384  
 Phe Pro Arg Leu Pro Arg Asp Gln Leu Tyr Arg Pro Pro Ser Asn Gly  
 115 120 125  
 50 gtg gaa gcc cct gag gag tcc ttg gac ctt cct gca gag ctg aga cca 432  
 Val Glu Ala Pro Glu Glu Ser Leu Asp Leu Pro Ala Glu Leu Arg Pro  
 130 135 140  
 cat ggt ccc cag gct cca tcc cta gct gcc gtg cct aga ccc cct agc 480  
 His Gly Pro Gln Ala Pro Ser Leu Ala Ala Val Pro Arg Pro Pro Ser  
 145 150 155 160  
 55 aac ccc tta gcc cga gga act cta aga acc agc aat ttg cca gaa gaa 528  
 Asn Pro Leu Ala Arg Gly Thr Leu Arg Thr Ser Asn Leu Pro Glu Glu  
 165 170 175

tta cgg aaa gtc ttt atc act tat tct atg gac aca gcc atg gag gtg 576  
 Leu Arg Lys Val Phe Ile Thr Tyr Ser Met Asp Thr Ala Met Glu Val  
 180 185 190

5     gtg aaa ttt gtg aac ttt ctg ttg gtg aac ggc ttc caa act gcg att 624  
 Val Lys Phe Val Asn Phe Leu Leu Val Asn Gly Phe Gln Thr Ala Ile  
 195 200 205

10     gac ata ttt gag gat aga atc cgg ggt att gat atc att aaa tgg atg 672  
 Asp Ile Phe Glu Asp Arg Ile Arg Gly Ile Asp Ile Ile Lys Trp Met  
 210 215 220

15     gag cgc tat ctt cga gat aag aca gtg atg ata atc gta gca atc agc 720  
 Glu Arg Tyr Leu Arg Asp Lys Thr Val Met Ile Ile Val Ala Ile Ser  
 225 230 235 240

20     ccc aaa tac aaa cag gat gtg gaa ggc gct gag tgc cag ctg gac gag 768  
 Pro Lys Tyr Lys Gln Asp Val Glu Gly Ala Glu Ser Gln Leu Asp Glu  
 245 250 255

25     gac gag cat ggc tta cat act aag tac att cat cgg atg atg cag att 816  
 Asp Glu His Gly Leu His Thr Lys Tyr Ile His Arg Met Met Gln Ile  
 260 265 270

30     gag ttc ata agt cag gga agc atg aac ttc aga ttc atc cct gtg ctc 864  
 Glu Phe Ile Ser Gln Gly Ser Met Asn Phe Arg Phe Ile Pro Val Leu  
 275 280 285

35     ttc cca aat gcc aag aag gag cat gtg ccg acc tgg ctt cag aac act 912  
 Phe Pro Asn Ala Lys Lys Glu His Val Pro Thr Trp Leu Gln Asn Thr  
 290 295 300

40     cat gtt tac agc tgg ccc aag aat aag aaa aac atc ctg ctg cgg ctg 960  
 His Val Tyr Ser Trp Pro Lys Asn Lys Lys Asn Ile Leu Leu Arg Leu  
 305 310 315 320

45     ctc agg gag gaa gag tat gtg gct cct ccc cga ggc cct ctg ccc acc 1008  
 Leu Arg Glu Glu Glu Tyr Val Ala Pro Pro Arg Gly Pro Leu Pro Thr  
 325 330 335

50     ctt cag gtg gta ccc ttg tgacgatggc cactccagct cagtgccagc 1056  
 Leu Gln Val Val Pro Leu  
 340

55     ctgtttctcac agcattcttc tagcggagct ggctggtggc acccaggccc tggaacacct 1116  
 cttctacaga gtcctctgtc tcttgagtct gagttgtcct cgctgggctt ccagagcttc 1176  
 agtgcctgga tgctgcaggt gacagaaaca aacatctatg accacaaaaa ctctcatcac 1236  
 ttcagctact tttatgagtc ggtcagatgc tctgtgtcct tagaccagtc taaatcatgc 1296  
 tcaaataata aaatgattat tcttttgt 1323

55     QDLPGPLRSRELPPQFELERYPMNAQLLPPHPSPQAPWNCQYYCPGGPYHHQVPHGHGYPPAAAYQQVLQPA  
 LPGQVLPGARARGPRPVQKVLNDSSPDQEERPAQRDFSFPRLPRDQLYRPPSNGVEAPEESLDLPAELRP  
 HGPQAPSLAAVPRPPSNPLARGTLRTSNLPEELRKVFITYSMDTAMEVVVKFVNFLLVNGFQTATDIFEDRIR  
 GIDIWKWMERYLRDKTVMIIVAISPKYKQDVEGAESQLDEDEHGLHTKYIHRMMQIEFISQGSMMNFRFIPVL  
 FPNACKKEHVPTWLQNTHVYSWPKNKNILLRLLREEEYVAPPRGPLPTLQVVPL.

Reverse translation of rodent, e.g., mouse, DCRS6 (SEQ ID NO: 27):

```

5   cargayytnc cnggnccnyt nmgnwsnmgn garytnccnc cncarttyga rytngarmgn 60
   tayccnatga aygcncaryt nytnccnccn cayccnwsnc cncargcncc ntggaaytgy 120
10  cartaytayt gyccngggngg nccntaycay caycargtnc cncayggncn yggntayccn 180
   ccngcngcng cntaycarca rgtnytncar ccngcnytn cnggncargt nytnccnggn 240
   gcnmgngcnm gnggnccnmg nccngtn car aargtnathy tnaaygayws nwsnccncar 300
15  gaycargarg armgnccngc ncarmnggay ttywsnttyc cnmgnytncc nmnggaycar 360
   ytnntaymgnc cncnwsnaa yggngtngar gcncngarg arwsnytn ga yytnccngcn 420
   garytnmgnc cncayggnc nccargcnccn wsnytnngcng cngtnccnmg nccnccnwsn 480
20  aayccnytn cnmngngnac nytnmgna cn wsnaayytnc cngargaryt nmgnaargtn 540
   ttyathacnt aywsnatgga yacngcnatg gargtngtna arttygtna yttyytnytn 600
25  gtnaayggnt tyacaracngc nathgayath ttygargaym gnathmgngg nathgayath 660
   athaartgga tggarmgnta yytnmgngay aaracngtna tgathathgt ngcnathwsn 720
   ccnaartaya arcargaygt ngarggngcn garwsncary tngaygarga ygarcaaygg 780
30  ytncaayacna artayathca ymgntatgat carathgart tyathwsnca rggnwsnatg 840
   aayttymgnt tyathccngt nytnnttyccn aaygcnaara argarcaygt nccnacntgg 900
35  ytncaraaya cncaygtnta ywsntggccn aaraayaara araayathyt nytnmgnytn 960
   ytnmgngarg argartaygt ngcnccnccn mgnggncny tnccnacnytn ncargtngtn 1020
   ccnytn 1026
40

```

Table 6: Alignment of the cytoplasmic portions of various cytokine receptor subunits. The IL-17R\_Hu (SEQ ID NO: 28) is GenBank AAB99730.1(U58917), gi|7657230; the IL-17R\_Mu (SEQ ID NO: 29) is GenBank AAC52357.1(U31993), gi|6680411; the IL-17R\_Ce (SEQ ID NO: 30) is GenBank AAA811100.1(U39997), gi|1353171; and the DCRS6\_Ce (SEQ ID NO: 31) is EMBCAA90543.1(Z50177), gi|7503597. Of particular interest are motifs or features corresponding, in primate DCRS8 to: R/K at 339/340; D/E at 348/349; alpha helical regions from H353-Q365, C370-S381, E389-H396, K410-D414, and D485-H495; beta sheet regions correspond to F400-V404 and F458-Y462; E at 431; E/D at 442/443; Y/F at 458; D/E at 468-470; Y/F at 481; and Q/R/F at 523.

45  
50

5	DCRS7_Mu DCRS7_Hu IL-17R_Hu IL-17R_Mu DCRS10 DCRS10_Mu DCRS9_Hu DCRS8_Hu IL-17R_Ce DCRS6_Hu DCRS6_Ce	RTALLLSADG-AGYERLVGALASALSQMP---LRVAVDLWSRRE-LSAHGALAWFHHQR RAALLLYSADD-SGFERLVGALASALCQLP---LRVAVDLWSRRE-LSAQGPVAVFHAQR RKVWIIYSADH-PLYVDVVLKFAQFLLTACG--TEVALDLLLEEQA-ISEAGVMTWVGRQK RKVWIVYSADH-PLYVEVVVLKFAQFLITACG--TEVALDLLEEQV-ISEVGVMTWVSRQK RKVFITYSMD---TAMEVVVKFVNFLVNG---FQTAIDIFEDR--IRGIDI IKWMERYL RKVFITYSMD---TAMEVVVKFVNFLVNG---FQTAIDIFEDR--IRGIDI IKWMERYL RPVLLLHAADS-EAQRRLVGALAE LLRAALGGGRDVIDLWEGRH-VARVGPLPWLWAAR PKVFLCYSSKDGQNHMNVVQCFAFYFLQDFCG--CEVALDLWEDFS-LCREGQREWVIQKI VKVMIVYADDN-DLHTDCVKKLVENLRNCAS--CDPVFDLEKLI--TAEIVPSRWLVQDI IKVLVVPSEI--CFHHTICYFTEFLQNHCR--SEVILEKWQKKK-IAEMGPVQWLATQK FKVMLVCPEVS-GRDEDFMMRIADALKKSN--NKVVCDRWFEFSKNAEENMLHWVYEQT .: . : : *
15	DCRS7_Mu DCRS7_Hu IL-17R_Hu IL-17R_Mu DCRS10 DCRS10_Mu DCRS9_Hu DCRS8_Hu IL-17R_Ce DCRS6_Hu DCRS6_Ce	RRILQEGGVVILLFSPAAVAQCQ---QWLQLQTVEP---GP---HDA LA A W L S C V L P D F L RQTLQEGGVVLLFSPGAVALCS---EWLQDGVSGPGAHP---HDAFRASLSCVLPDFL QEMVESNSKIIIVLCSRGTRAKWQALLGRGAP-VRLRCDHGKPV-GDLFTAAMNMILPDFK QEMVESNSKIIILCSRGTOAKWKAILGWAEPVQLRCDHWKPA-GDLFTAAMNMILPDFK R---DKTVMIIIVAI SPKYQDVE---GAESQLDED-EHGL---HTKYIHRM-MQIEFIK R---DKTVMIIIVAI SPKYQDVE---GAESQLDED-EHGL---HTKYIHRM-MQIEFIS TRVAREQGTVLLWSGADLRPV S---GPD P-RAAP-----LLA---LLHAAP H---ESQFIIIVCSKGMKYFVD---KKNYKHKGGRGSGK---GELFLVAVSAIAEKL R S---SLKKFIIIVSDCAEKILD---TEASETHQLVQARP--FADLFGPAMEMIIRDAT K---AADKVVFLLSNDVNSVCD---GTCGKSEGSPSENS---QDLFPLAFNLFCSDLR K---IAEKIIIVFHSAYYHPRCG---IYDVINNFPCTDPR-----LAHIALT---PEAQ .: . *
30	DCRS7_Mu DCRS7_Hu IL-17R_Hu IL-17R_Mu DCRS10 DCRS10_Mu DCRS9_Hu DCRS8_Hu IL-17R_Ce DCRS6_Hu DCRS6_Ce	QGRATGR-----YGVVYFDGLLHPDSVPSPFRVAPLFSLP-SQLPAFLDALQ--GGCSTS QGRAPGS-----YVGACFDRLLHPDAVPALFRTVPVFTLP-SQLPDLFGALQ--QPRAPR RPACFGT-----YVVCYFSEVSCDGDVPDLFGAAPRYPLM-DRFEEVYFRIQ--DLEMFO RPACFGT-----YVVCYFSGICSERDVPDLFNITSRYPLM-DRFEEVYFRIQ--DLEMFE QGS MNFR-----FIPVLFPNAK-KEHVPTWLQNTHVYSWP-KNKKNILLRLL-REEEYVA QGS MNFR-----FIPVLFPNAK-KEHVPTWLQNTHVYSWP-KNKKNILLRLL-REEEYVA RPL-----LLLAYFSRLCAKGDIPPLRALPRYRL- RDLPRLLRALD--ARPFAE QAKQSSSAALS KFI AVYFDYSC-EGDVP GILD LSTKYRLM-DNLPQLCSHLHSRDHGLQE HNFPEAR---KKYAVVRFNYS P--HVPPNLAILNLPTFIPEQFAQLTAF LHN-VEHTER SQIHLHK----YVVVYFREID-TKDDYNALSVC PKYHLM-KDATAFCAELL---HVKQQ RSVPKEV----EYVLPRDQKLL--EDAFDITIADPLVIDIPIEDVAIPENVP--IH HES C
40	DCRS7_Mu DCRS7_Hu IL-17R_Hu IL-17R_Mu DCRS10 DCRS10_Mu DCRS9_Hu DCRS8_Hu IL-17R_Ce DCRS6_Hu DCRS6_Ce	AGRPAD RVER-----VT-----QALRSALDSCTS----- SGRLQERAEQ-----VS-----RALQPALDSYFHPP----- PGRMHRVGELSGDNYLRS---PGRQLRAALDRFRDWQVRCPDW PGRMHVREL TGDNYLQS---PSGRQLKEAVLRFQEWQTQCPDW P----PRGPL-----PTLQVVPL----- P----PRGPL-----PTLQVVPL----- ATSWGRLGAR-----QRRQSRLELCSR----- PGQHTRQGS R-----RNYFRSKSGRSLYVAICNMHQFIDE EPDW ANVTQNISEA-----Q-----IHEWNLCASRMMSFFVRNPNW VS----AGKR-----SQACHDGCCSL----- DSIDSRNNSK-----THSTDSGVSSLS S-----NS--

Table 6 shows comparison of the available sequences of primate, rodent, and various other receptors. Various conserved residues are aligned and indicated. The structurally homologous cytoplasmic domains most likely signal through pathways like IL-17, e.g., through NFkB. Similar to IL-1 signalling, it is likely that these receptors are involved in innate immunity and/or development.

As used herein, the term DCRS shall be used to describe a protein comprising amino acid sequences shown in Tables 1-5, respectively. In many cases, a substantial fragment thereof will be functionally or structurally equivalent, including, e.g., an extracellular or intracellular domain. The invention also includes a protein variation of the respective DCRS allele whose sequence is provided, e.g., a mutein or soluble extracellular construct. Typically, such agonists or antagonists will exhibit less than about 10% sequence differences, and thus will often have between 1 and 11 substitutions, e.g., 2-, 3-, 5-, 7-fold, and others. It also encompasses allelic and other variants, e.g., natural polymorphic, of the protein described. Typically, it will bind to its corresponding biological ligand, perhaps in a dimerized state with an alpha receptor subunit, with high affinity, e.g., at least about 100 nM, usually better than about 30 nM, preferably better than about 10 nM, and more preferably at better than about 3 nM. The term shall also be used herein to refer to related naturally occurring forms, e.g., alleles, polymorphic variants, and metabolic variants of the mammalian protein. Preferred forms of the receptor complexes will bind the appropriate ligand with an affinity and selectivity appropriate for a ligand-receptor interaction.

This invention also encompasses combinations of proteins or peptides having substantial amino acid sequence identity with an amino acid sequence in Tables 1-5. It will include sequence variants with relatively few residue substitutions, e.g., preferably less than about 3-5.

A substantial polypeptide "fragment", or "segment", is a stretch of amino acid residues of at least about 8 amino acids, generally at least 10 amino acids, more generally at least 12 amino acids, often at least 14 amino acids, more often at least 16 amino acids, typically at least 18 amino acids, more typically at least 20 amino acids, usually at least 22 amino acids, more usually at least 24 amino acids, preferably at least 26 amino acids, more preferably at least 28 amino acids, and, in particularly preferred embodiments, at least about 30 or more amino acids. This includes, e.g., 40, 50, 60, 70, 85, 100, 115, 130, 150, and other lengths. Sequences of segments of different proteins can be compared to one another over appropriate length stretches, typically between conserved motifs. In many situations, fragments may exhibit functional properties of the intact subunits, e.g., the extracellular domain of the transmembrane receptor may retain the ligand binding features, and may be used to prepare a soluble receptor-like complex.



Amino acid sequence homology, or sequence identity, is determined by optimizing residue matches. In some comparisons, gaps may be introduced, as required. See, e.g., Needleham, et al., (1970) *J. Mol. Biol.* 48:443-453; Sankoff, et al., (1983) chapter one in Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison, Addison-Wesley, Reading, MA; and software packages from IntelliGenetics, Mountain View, CA; and the University of Wisconsin Genetics Computer Group (GCG), Madison, WI; each of which is incorporated herein by reference. This changes when considering conservative substitutions as matches. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid; asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. Homologous amino acid sequences are intended to include natural allelic and interspecies variations in the cytokine sequence. Typical homologous proteins or peptides will have from 50-100% homology (if gaps can be introduced), to 60-100% homology (if conservative substitutions are included) with an amino acid sequence segment of, e.g., Table 3 or 4. Homology measures will be at least about 70%, generally at least 76%, more generally at least 81%, often at least 85%, more often at least 88%, typically at least 90%, more typically at least 92%, usually at least 94%, more usually at least 95%, preferably at least 96%, and more preferably at least 97%, and in particularly preferred embodiments, at least 98% or more. The degree of homology will vary with the length of the compared segments. Homologous proteins or peptides, such as the allelic variants, will share most biological activities with the embodiments described in Tables 1-5.

As used herein, the term "biological activity" is used to describe, without limitation, effects on inflammatory responses, innate immunity, and/or morphogenic development by cytokine-like ligands. For example, these receptors should mediate phosphatase or phosphorylase activities, which activities are easily measured by standard procedures. See, e.g., Hardie, et al. (eds. 1995) The Protein Kinase FactBook vols. I and II, Academic Press, San Diego, CA; Hanks, et al. (1991) Meth. Enzymol. 200:38-62; Hunter, et al. (1992) Cell 70:375-388; Lewin (1990) Cell 61:743-752; Pines, et al. (1991) Cold Spring Harbor Symp. Quant. Biol. 56:449-463; and Parker, et al. (1993) Nature 363:736-738. The receptors, or portions thereof, may be useful as phosphate labeling enzymes to label general or specific substrates. The subunits may also be functional immunogens to elicit recognizing antibodies, or antigens capable of binding antibodies.

The terms ligand, agonist, antagonist, and analog of, e.g., a DCRS8 or DCRS9, include molecules that modulate the characteristic cellular responses to cytokine ligand proteins, as well as molecules possessing the more standard structural binding competition features of ligand-receptor interactions, e.g., where the receptor is a natural

receptor or an antibody. The cellular responses likely are typically mediated through receptor tyrosine kinase pathways.

Also, a ligand is a molecule which serves either as a natural ligand to which said receptor, or an analog thereof, binds, or a molecule which is a functional analog of the natural ligand. The functional analog may be a ligand with structural modifications, or may be a wholly unrelated molecule which has a molecular shape which interacts with the appropriate ligand binding determinants. The ligands may serve as agonists or antagonists, see, e.g., Goodman, et al. (eds. 1990) Goodman & Gilman's: The Pharmacological Bases of Therapeutics, Pergamon Press, New York.

Rational drug design may also be based upon structural studies of the molecular shapes of a receptor or antibody and other effectors or ligands. See, e.g., Herz, et al. (1997) J. Recept. Signal Transduct. Res. 17:671-776; and Chaiken, et al. (1996) Trends Biotechnol. 14:369-375. Effectors may be other proteins which mediate other functions in response to ligand binding, or other proteins which normally interact with the receptor. One means for determining which sites interact with specific other proteins is a physical structure determination, e.g., x-ray crystallography or 2 dimensional NMR techniques. These will provide guidance as to which amino acid residues form molecular contact regions. For a detailed description of protein structural determination, see, e.g., Blundell and Johnson (1976) Protein Crystallography, Academic Press, New York, which is hereby incorporated herein by reference.

## II. Activities

The cytokine receptor-like proteins will have a number of different biological activities, e.g., modulating cell proliferation, or in phosphate metabolism, being added to or removed from specific substrates, typically proteins. Such will generally result in modulation of an inflammatory function, other innate immunity response, or a morphological effect. The subunit will probably have a specific low affinity binding to the ligand.

The DCRS8 and DCRS9 have characteristic motifs of receptors signaling through the JAK pathway. See, e.g., Ihle, et al. (1997) Stem Cells 15(suppl. 1):105-111; Silvennoinen, et al. (1997) APMIS 105:497-509; Levy (1997) Cytokine Growth Factor Review 8:81-90; Winston and Hunter (1996) Current Biol. 6:668-671; Barrett (1996) Baillieres Clin. Gastroenterol. 10:1-15; and Briscoe, et al. (1996) Philos. Trans. R. Soc. Lond. B. Biol. Sci. 351:167-171.

The biological activities of the cytokine receptor subunits will be related to addition or removal of phosphate moieties to substrates, typically in a specific manner, but occasionally in a non specific manner. Substrates may be identified, or conditions for

enzymatic activity may be assayed by standard methods, e.g., as described in Hardie, et al. (eds. 1995) The Protein Kinase FactBook vols. I and II, Academic Press, San Diego, CA; Hanks, et al. (1991) Meth. Enzymol. 200:38-62; Hunter, et al. (1992) Cell 70:375-388; Lewin (1990) Cell 61:743-752; Pines, et al. (1991) Cold Spring Harbor Symp. Quant. Biol. 56:449-463; and Parker, et al. (1993) Nature 363:736-738.

The receptor subunits may combine to form functional complexes, e.g., which may be useful for binding ligand or preparing antibodies. These will have substantial diagnostic uses, including detection or quantitation.

### 10 III. Nucleic Acids

This invention contemplates use of isolated nucleic acid or fragments, e.g., which encode these or closely related proteins, or fragments thereof, e.g., to encode a corresponding polypeptide, preferably one which is biologically active. In addition, this invention covers isolated or recombinant DNAs which encode combinations of such  
15 proteins or polypeptides having characteristic sequences, e.g., of the DCRSs. Typically, the nucleic acid is capable of hybridizing, under appropriate conditions, with a nucleic acid sequence segment shown in Tables 1-5, but preferably not with a corresponding segment of other receptors described in Table 6. Said biologically active protein or polypeptide can be a full length protein, or fragment, and will typically have a segment of  
20 amino acid sequence highly homologous, e.g., exhibiting significant stretches of identity, to one shown in Tables 1-5. Further, this invention covers the use of isolated or recombinant nucleic acid, or fragments thereof, which encode proteins having fragments which are equivalent to the DCRS8 or DCRS9 proteins. The isolated nucleic acids can have the respective regulatory sequences in the 5' and 3' flanks, e.g., promoters,  
25 enhancers, poly-A addition signals, and others from the natural gene. Combinations, as described, are also provided.

An "isolated" nucleic acid is a nucleic acid, e.g., an RNA, DNA, or a mixed polymer, which is substantially pure, e.g., separated from other components which naturally accompany a native sequence, such as ribosomes, polymerases, and flanking  
30 genomic sequences from the originating species. The term embraces a nucleic acid sequence which has been removed from its naturally occurring environment, and includes recombinant or cloned DNA isolates, which are thereby distinguishable from naturally occurring compositions, and chemically synthesized analogs or analogs biologically synthesized by heterologous systems. A substantially pure molecule includes isolated  
35 forms of the molecule, either completely or substantially pure.

An isolated nucleic acid will generally be a homogeneous composition of molecules, but will, in some embodiments, contain heterogeneity, preferably minor. This

heterogeneity is typically found at the polymer ends or portions not critical to a desired biological function or activity.

A "recombinant" nucleic acid is typically defined either by its method of production or its structure. In reference to its method of production, e.g., a product made by a process, the process is use of recombinant nucleic acid techniques, e.g., involving human intervention in the nucleotide sequence. Typically this intervention involves in vitro manipulation, although under certain circumstances it may involve more classical animal breeding techniques. Alternatively, it can be a nucleic acid made by generating a sequence comprising fusion of two fragments which are not naturally contiguous to each other, but is meant to exclude products of nature, e.g., naturally occurring mutants as found in their natural state. Thus, for example, products made by transforming cells with an unnaturally occurring vector is encompassed, as are nucleic acids comprising sequence derived using any synthetic oligonucleotide process. Such a process is often done to replace a codon with a redundant codon encoding the same or a conservative amino acid, while typically introducing or removing a restriction enzyme sequence recognition site. Alternatively, the process is performed to join together nucleic acid segments of desired functions to generate a single genetic entity comprising a desired combination of functions not found in the commonly available natural forms, e.g., encoding a fusion protein. Restriction enzyme recognition sites are often the target of such artificial manipulations, but other site specific targets, e.g., promoters, DNA replication sites, regulation sequences, control sequences, or other useful features may be incorporated by design. A similar concept is intended for a recombinant, e.g., fusion, polypeptide. This will include a dimeric repeat. Specifically included are synthetic nucleic acids which, by genetic code redundancy, encode equivalent polypeptides to fragments of DCRSs and fusions of sequences from various different related molecules, e.g., other cytokine receptor family members.

A "fragment" in a nucleic acid context is a contiguous segment of at least about 17 nucleotides, generally at least 21 nucleotides, more generally at least 25 nucleotides, ordinarily at least 30 nucleotides, more ordinarily at least 35 nucleotides, often at least 39 nucleotides, more often at least 45 nucleotides, typically at least 50 nucleotides, more typically at least 55 nucleotides, usually at least 60 nucleotides, more usually at least 66 nucleotides, preferably at least 72 nucleotides, more preferably at least 79 nucleotides, and in particularly preferred embodiments will be at least 85 or more nucleotides. Typically, fragments of different genetic sequences can be compared to one another over appropriate length stretches, particularly defined segments such as the domains described below.

5 A nucleic acid which codes for the DCRS8 or DCRS9 will be particularly useful to identify genes, mRNA, and cDNA species which code for itself or closely related proteins, as well as DNAs which code for polymorphic, allelic, or other genetic variants, e.g., from different individuals or related species. Preferred probes for such screens are those regions of the interleukin which are conserved between different polymorphic variants or which contain nucleotides which lack specificity, and will preferably be full length or nearly so. In other situations, polymorphic variant specific sequences will be more useful.

10 This invention further covers recombinant nucleic acid molecules and fragments having a nucleic acid sequence identical to or highly homologous to the isolated DNA set forth herein. In particular, the sequences will often be operably linked to DNA segments which control transcription, translation, and DNA replication. These additional segments typically assist in expression of the desired nucleic acid segment.

15 Homologous, or highly identical, nucleic acid sequences, when compared to one another, e.g., DCRS8 sequences, exhibit significant similarity. The standards for homology in nucleic acids are either measures for homology generally used in the art by sequence comparison or based upon hybridization conditions. Comparative hybridization conditions are described in greater detail below.

20 Substantial identity in the nucleic acid sequence comparison context means either that the segments, or their complementary strands, when compared, are identical when optimally aligned, with appropriate nucleotide insertions or deletions, in at least about 60% of the nucleotides, generally at least 66%, ordinarily at least 71%, often at least 76%, more often at least 80%, usually at least 84%, more usually at least 88%, typically at least 91%, more typically at least about 93%, preferably at least about 95%, more preferably at least about 96 to 98% or more, and in particular embodiments, as high as about 99% or more of the nucleotides, including, e.g., segments encoding structural domains such as the segments described below. Alternatively, substantial identity will exist when the segments will hybridize under selective hybridization conditions, to a strand or its complement, typically using a sequence derived from Tables 1-5. Typically, selective hybridization will occur when there is at least about 55% homology over a stretch of at least about 14 nucleotides, more typically at least about 65%, preferably at least about 75%, and more preferably at least about 90%. See, Kanehisa (1984) Nucl. Acids Res. 12:203-213, which is incorporated herein by reference. The length of homology comparison, as described, may be over longer stretches, and in certain embodiments will be over a stretch of at least about 17 nucleotides, generally at least about 20 nucleotides, ordinarily at least about 24 nucleotides, usually at least about 28 nucleotides, typically at least about 32 nucleotides, more typically at least about 40 nucleotides, preferably at least

about 50 nucleotides, and more preferably at least about 75 to 100 or more nucleotides. This includes, e.g., 125, 150, 175, 200, 225, 246, 273, and other lengths.

Stringent conditions, in referring to homology in the hybridization context, will be stringent combined conditions of salt, temperature, organic solvents, and other parameters typically controlled in hybridization reactions. Stringent temperature conditions will usually include temperatures in excess of about 30 C, more usually in excess of about 37 C, typically in excess of about 45 C, more typically in excess of about 55 C, preferably in excess of about 65 C, and more preferably in excess of about 70 C. Stringent salt conditions will ordinarily be less than about 500 mM, usually less than about 400 mM, more usually less than about 300 mM, typically less than about 200 mM, preferably less than about 100 mM, and more preferably less than about 80 mM, even down to less than about 20 mM. However, the combination of parameters is much more important than the measure of any single parameter. See, e.g., Wetmur and Davidson (1968) J. Mol. Biol. 31:349-370, which is hereby incorporated herein by reference.

The isolated DNA can be readily modified by nucleotide substitutions, nucleotide deletions, nucleotide insertions, and inversions of nucleotide stretches. These modifications result in novel DNA sequences which encode this protein or its derivatives. These modified sequences can be used to produce mutant proteins (muteins) or to enhance the expression of variant species. Enhanced expression may involve gene amplification, increased transcription, increased translation, and other mechanisms. Such mutant DCRS8-like derivatives include predetermined or site-specific mutations of the protein or its fragments, including silent mutations using genetic code degeneracy. "Mutant DCRS8" as used herein encompasses a polypeptide otherwise falling within the homology definition of the DCRS8 as set forth above, but having an amino acid sequence which differs from that of other cytokine receptor-like proteins as found in nature, whether by way of deletion, substitution, or insertion. In particular, "site specific mutant DCRS8" encompasses a protein having substantial sequence identity with a protein of Table 3, and typically shares most of the biological activities or effects of the forms disclosed herein.

Although site specific mutation sites are predetermined, mutants need not be site specific. Mammalian DCRS8 mutagenesis can be achieved by making amino acid insertions or deletions in the gene, coupled with expression. Substitutions, deletions, insertions, or many combinations may be generated to arrive at a final construct. Insertions include amino- or carboxy- terminal fusions. Random mutagenesis can be conducted at a target codon and the expressed mammalian DCRS mutants can then be screened for the desired activity, providing some aspect of a structure-activity relationship. Methods for making substitution mutations at predetermined sites in DNA

having a known sequence are well known in the art, e.g., by M13 primer mutagenesis. See also Sambrook, et al. (1989) and Ausubel, et al. (1987 and periodic Supplements).

5 The mutations in the DNA normally should not place coding sequences out of reading frames and preferably will not create complementary regions that could hybridize to produce secondary mRNA structure such as loops or hairpins.

The phosphoramidite method described by Beaucage and Carruthers (1981) Tetra. Letts. 22:1859-1862, will produce suitable synthetic DNA fragments. A double stranded fragment will often be obtained either by synthesizing the complementary strand and annealing the strand together under appropriate conditions or by adding the  
10 complementary strand using DNA polymerase with an appropriate primer sequence.

Polymerase chain reaction (PCR) techniques can often be applied in mutagenesis. Alternatively, mutagenesis primers are commonly used methods for generating defined mutations at predetermined sites. See, e.g., Innis, et al. (eds. 1990) PCR Protocols: A Guide to Methods and Applications Academic Press, San Diego, CA; and Dieffenbach  
15 and Dveksler (1995; eds.) PCR Primer: A Laboratory Manual Cold Spring Harbor Press, CSH, NY.

Certain embodiments of the invention are directed to combination compositions comprising the receptor or ligand sequences described. In other embodiments, functional portions of the sequences may be joined to encode fusion proteins. In other forms,  
20 variants of the described sequences may be substituted.

#### IV. Proteins, Peptides

As described above, the present invention encompasses primate DCRS6-10, e.g., whose sequences are disclosed in Tables 1-5, and described above. Allelic and other  
25 variants are also contemplated, including, e.g., fusion proteins combining portions of such sequences with others, including, e.g., epitope tags and functional domains.

The present invention also provides recombinant proteins, e.g., heterologous fusion proteins using segments from these primate or rodent proteins. A heterologous fusion protein is a fusion of proteins or segments which are naturally not normally fused  
30 in the same manner. Thus, the fusion product of, e.g., a DCRS8 with another cytokine receptor is a continuous protein molecule having sequences fused in a typical peptide linkage, typically made as a single translation product and exhibiting properties, e.g., sequence or antigenicity, derived from each source peptide. A similar concept applies to heterologous nucleic acid sequences. Combinations of various designated proteins into  
35 complexes are also provided.

In addition, new constructs may be made from combining similar functional or structural domains from other related proteins, e.g., cytokine receptors or Toll-like

receptors, including species variants. For example, ligand-binding or other segments may be "swapped" between different new fusion polypeptides or fragments. See, e.g., Cunningham, et al. (1989) Science 243:1330-1336; and O'Dowd, et al. (1988) J. Biol. Chem. 263:15985-15992, each of which is incorporated herein by reference. Thus, new chimeric polypeptides exhibiting new combinations of specificities will result from the functional linkage of receptor-binding specificities. For example, the ligand binding domains from other related receptor molecules may be added or substituted for other domains of this or related proteins. The resulting protein will often have hybrid function and properties. For example, a fusion protein may include a targeting domain which may serve to provide sequestering of the fusion protein to a particular subcellular organelle.

Candidate fusion partners and sequences can be selected from various sequence data bases, e.g., GenBank, c/o IntelliGenetics, Mountain View, CA; and BCG, University of Wisconsin Biotechnology Computing Group, Madison, WI, which are each incorporated herein by reference. In particular, combinations of polypeptide sequences provided in Tables 1-5 are particularly preferred. Variant forms of the proteins may be substituted in the described combinations.

The present invention particularly provides muteins which bind cytokine-like ligands, and/or which are affected in signal transduction. Structural alignment of human DCRSs with other members of the cytokine receptor family show conserved features/residues. See Table 6. Alignment of the human DCRS8 sequence with other members of the cytokine receptor family indicates various structural and functionally shared features. See also, Bazan, et al. (1996) Nature 379:591; Lodi, et al. (1994) Science 263:1762-1766; Sayle and Milner-White (1995) TIBS 20:374-376; and Gronenberg, et al. (1991) Protein Engineering 4:263-269.

Substitutions with either mouse sequences or human sequences are particularly preferred. Conversely, conservative substitutions away from the ligand binding interaction regions will probably preserve most signaling activities; and conservative substitutions away from the intracellular domains will probably preserve most ligand binding properties.

"Derivatives" of the primate DCRS8 include amino acid sequence mutants, glycosylation variants, metabolic derivatives and covalent or aggregative conjugates with other chemical moieties. Covalent derivatives can be prepared by linkage of functionalities to groups which are found in the DCRS8 amino acid side chains or at the N- or C- termini, e.g., by means which are well known in the art. These derivatives can include, without limitation, aliphatic esters or amides of the carboxyl terminus, or of residues containing carboxyl side chains, O-acyl derivatives of hydroxyl group-containing residues, and N-acyl derivatives of the amino terminal amino acid or amino-group



containing residues, e.g., lysine or arginine. Acyl groups are selected from the group of alkyl-moieties, including C3 to C18 normal alkyl, thereby forming alkanoyl aroyl species.

5 In particular, glycosylation alterations are included, e.g., made by modifying the glycosylation patterns of a polypeptide during its synthesis and processing, or in further processing steps. Particularly preferred means for accomplishing this are by exposing the polypeptide to glycosylating enzymes derived from cells which normally provide such processing, e.g., mammalian glycosylation enzymes. Deglycosylation enzymes are also contemplated. Also embraced are versions of the same primary amino acid sequence which have other minor modifications, including phosphorylated amino acid residues,  
10 e.g., phosphotyrosine, phosphoserine, or phosphothreonine.

A major group of derivatives are covalent conjugates of the receptors or fragments thereof with other proteins of polypeptides. These derivatives can be synthesized in recombinant culture such as N- or C-terminal fusions or by the use of agents known in the art for their usefulness in cross-linking proteins through reactive side groups. Preferred  
15 derivatization sites with cross-linking agents are at free amino groups, carbohydrate moieties, and cysteine residues.

Fusion polypeptides between the receptors and other homologous or heterologous proteins are also provided. Homologous polypeptides may be fusions between different receptors, resulting in, for instance, a hybrid protein exhibiting binding specificity for  
20 multiple different cytokine ligands, or a receptor which may have broadened or weakened specificity of substrate effect. Likewise, heterologous fusions may be constructed which would exhibit a combination of properties or activities of the derivative proteins. Typical examples are fusions of a reporter polypeptide, e.g., luciferase, with a segment or domain of a receptor, e.g., a ligand-binding segment, so that the presence or location of a desired  
25 ligand may be easily determined. See, e.g., Dull, et al., U.S. Patent No. 4,859,609, which is hereby incorporated herein by reference. Other gene fusion partners include glutathione-S-transferase (GST), bacterial  $\beta$ -galactosidase, trpE, Protein A,  $\beta$ -lactamase, alpha amylase, alcohol dehydrogenase, and yeast alpha mating factor. See, e.g., Godowski, et al. (1988) Science 241:812-816. Labeled proteins will often be substituted  
30 in the described combinations of proteins.

The phosphoramidite method described by Beaucage and Carruthers (1981) Tetra. Letts. 22:1859-1862, will produce suitable synthetic DNA fragments. A double stranded fragment will often be obtained either by synthesizing the complementary strand and annealing the strand together under appropriate conditions or by adding the  
35 complementary strand using DNA polymerase with an appropriate primer sequence.

Such polypeptides may also have amino acid residues which have been chemically modified by phosphorylation, sulfonation, biotinylation, or the addition or removal of

other moieties, particularly those which have molecular shapes similar to phosphate groups. In some embodiments, the modifications will be useful labeling reagents, or serve as purification targets, e.g., affinity ligands.

5 Fusion proteins will typically be made by either recombinant nucleic acid methods or by synthetic polypeptide methods. Techniques for nucleic acid manipulation and expression are described generally, for example, in Sambrook, et al. (1989) Molecular Cloning: A Laboratory Manual (2d ed.), Vols. 1-3, Cold Spring Harbor Laboratory, and Ausubel, et al. (eds. 1987 and periodic supplements) Current Protocols in Molecular Biology, Greene/Wiley, New York, which are each incorporated herein by reference.  
10 Techniques for synthesis of polypeptides are described, for example, in Merrifield (1963) J. Amer. Chem. Soc. 85:2149-2156; Merrifield (1986) Science 232: 341-347; and Atherton, et al. (1989) Solid Phase Peptide Synthesis: A Practical Approach, IRL Press, Oxford; each of which is incorporated herein by reference. See also Dawson, et al. (1994) Science 266:776-779 for methods to make larger polypeptides.

15 This invention also contemplates the use of derivatives of a DCRS8 other than variations in amino acid sequence or glycosylation. Such derivatives may involve covalent or aggregative association with chemical moieties. These derivatives generally fall into three classes: (1) salts, (2) side chain and terminal residue covalent modifications, and (3) adsorption complexes, for example with cell membranes. Such covalent or  
20 aggregative derivatives are useful as immunogens, as reagents in immunoassays, or in purification methods such as for affinity purification of a receptor or other binding molecule, e.g., an antibody. For example, a cytokine ligand can be immobilized by covalent bonding to a solid support such as cyanogen bromide-activated Sepharose, by methods which are well known in the art, or adsorbed onto polyolefin surfaces, with or  
25 without glutaraldehyde cross-linking, for use in the assay or purification of a cytokine receptor, antibodies, or other similar molecules. The ligand can also be labeled with a detectable group, for example radioiodinated by the chloramine T procedure, covalently bound to rare earth chelates, or conjugated to another fluorescent moiety for use in diagnostic assays.

30 A combination, e.g., including a DCRS8, of this invention can be used as an immunogen for the production of antisera or antibodies specific, e.g., capable of distinguishing between other cytokine receptor family members, for the combinations described. The complexes can be used to screen monoclonal antibodies or antigen-binding fragments prepared by immunization with various forms of impure preparations  
35 containing the protein. In particular, the term "antibodies" also encompasses antigen binding fragments of natural antibodies, e.g., Fab, Fab2, Fv, etc. The purified DCRS8 can also be used as a reagent to detect antibodies generated in response to the presence of

elevated levels of expression, or immunological disorders which lead to antibody production to the endogenous receptor. Additionally, DCRS8 fragments may also serve as immunogens to produce the antibodies of the present invention, as described immediately below. For example, this invention contemplates antibodies having binding affinity to or being raised against the amino acid sequences shown in Tables 1-5, fragments thereof, or various homologous peptides. In particular, this invention contemplates antibodies having binding affinity to, or having been raised against, specific fragments which are predicted to be, or actually are, exposed at the exterior protein surface of the native DCRS8 or DCRS9. Complexes of combinations of proteins will also be useful, and antibody preparations thereto can be made.

The blocking of physiological response to the receptor ligands may result from the inhibition of binding of the ligand to the receptor, likely through competitive inhibition. Thus, in vitro assays of the present invention will often use antibodies or antigen binding segments of these antibodies, or fragments attached to solid phase substrates. These assays will also allow for the diagnostic determination of the effects of either ligand binding region mutations and modifications, or other mutations and modifications, e.g., which affect signaling or enzymatic function.

This invention also contemplates the use of competitive drug screening assays, e.g., where neutralizing antibodies to the receptor complexes or fragments compete with a test compound for binding to a ligand or other antibody. In this manner, the neutralizing antibodies or fragments can be used to detect the presence of a polypeptide which shares one or more binding sites to a receptor and can also be used to occupy binding sites on a receptor that might otherwise bind a ligand.

## V. Making Nucleic Acids and Protein

DNA which encodes the protein or fragments thereof can be obtained by chemical synthesis, screening cDNA libraries, or by screening genomic libraries prepared from a wide variety of cell lines or tissue samples. Natural sequences can be isolated using standard methods and the sequences provided herein, e.g., in Tables 1-5. Other species counterparts can be identified by hybridization techniques, or by various PCR techniques, combined with or by searching in sequence databases, e.g., GenBank.

This DNA can be expressed in a wide variety of host cells for the synthesis of a full-length receptor or fragments which can in turn, for example, be used to generate polyclonal or monoclonal antibodies; for binding studies; for construction and expression of modified ligand binding or kinase/phosphatase domains; and for structure/function studies. Variants or fragments can be expressed in host cells that are transformed or transfected with appropriate expression vectors. These molecules can be substantially

free of protein or cellular contaminants, other than those derived from the recombinant host, and therefore are particularly useful in pharmaceutical compositions when combined with a pharmaceutically acceptable carrier and/or diluent. The protein, or portions thereof, may be expressed as fusions with other proteins. Combinations of the described  
5 proteins, or nucleic acids encoding them, are particularly interesting.

Expression vectors are typically self-replicating DNA or RNA constructs containing the desired receptor gene or its fragments, usually operably linked to suitable genetic control elements that are recognized in a suitable host cell. These control  
10 elements are capable of effecting expression within a suitable host. The multiple genes may be coordinately expressed, and may be on a polycistronic message. The specific type of control elements necessary to effect expression will depend upon the eventual host cell used. Generally, the genetic control elements can include a prokaryotic promoter system or a eukaryotic promoter expression control system, and typically include a transcriptional promoter, an optional operator to control the onset of transcription,  
15 transcription enhancers to elevate the level of mRNA expression, a sequence that encodes a suitable ribosome binding site, and sequences that terminate transcription and translation. Expression vectors also usually contain an origin of replication that allows the vector to replicate independently of the host cell.

The vectors of this invention include those which contain DNA which encodes a  
20 combination of proteins, as described, or a biologically active equivalent polypeptide. The DNA can be under the control of a viral promoter and can encode a selection marker. This invention further contemplates use of such expression vectors which are capable of expressing eukaryotic cDNAs coding for such proteins in a prokaryotic or eukaryotic host, where the vector is compatible with the host and where the eukaryotic cDNAs are  
25 inserted into the vector such that growth of the host containing the vector expresses the cDNAs in question. Usually, expression vectors are designed for stable replication in their host cells or for amplification to greatly increase the total number of copies of the desirable gene per cell. It is not always necessary to require that an expression vector replicate in a host cell, e.g., it is possible to effect transient expression of the protein or its  
30 fragments in various hosts using vectors that do not contain a replication origin that is recognized by the host cell. It is also possible to use vectors that cause integration of the protein encoding portions into the host DNA by recombination.

Vectors, as used herein, comprise plasmids, viruses, bacteriophage, integratable DNA fragments, and other vehicles which enable the integration of DNA fragments into  
35 the genome of the host. Expression vectors are specialized vectors which contain genetic control elements that effect expression of operably linked genes. Plasmids are the most commonly used form of vector but all other forms of vectors which serve an equivalent

function and which are, or become, known in the art are suitable for use herein. See, e.g., Pouwels, et al. (1985 and Supplements) Cloning Vectors: A Laboratory Manual, Elsevier, N.Y., and Rodriguez, et al. (eds. 1988) Vectors: A Survey of Molecular Cloning Vectors and Their Uses, Butterworth, Boston, which are incorporated herein by reference.

5 Transformed cells are cells, preferably mammalian, that have been transformed or transfected with vectors constructed using recombinant DNA techniques. Transformed host cells usually express the desired proteins, but for purposes of cloning, amplifying, and manipulating its DNA, do not need to express the subject proteins. This invention further contemplates culturing transformed cells in a nutrient medium, thus permitting the  
10 proteins to accumulate. The proteins can be recovered, either from the culture or, in certain instances, from the culture medium.

For purposes of this invention, nucleic sequences are operably linked when they are functionally related to each other. For example, DNA for a presequence or secretory leader is operably linked to a polypeptide if it is expressed as a preprotein or participates  
15 in directing the polypeptide to the cell membrane or in secretion of the polypeptide. A promoter is operably linked to a coding sequence if it controls the transcription of the polypeptide; a ribosome binding site is operably linked to a coding sequence if it is positioned to permit translation. Usually, operably linked means contiguous and in reading frame, however, certain genetic elements such as repressor genes are not  
20 contiguously linked but still bind to operator sequences that in turn control expression.

Suitable host cells include prokaryotes, lower eukaryotes, and higher eukaryotes. Prokaryotes include both gram negative and gram positive organisms, e.g., E. coli and B. subtilis. Lower eukaryotes include yeasts, e.g., S. cerevisiae and Pichia, and species of the genus Dictyostelium. Higher eukaryotes include established tissue culture cell lines  
25 from animal cells, both of non-mammalian origin, e.g., insect cells, and birds, and of mammalian origin, e.g., human, primates, and rodents.

Prokaryotic host-vector systems include a wide variety of vectors for many different species. As used herein, E. coli and its vectors will be used generically to include equivalent vectors used in other prokaryotes. A representative vector for  
30 amplifying DNA is pBR322 or many of its derivatives. Vectors that can be used to express the receptor or its fragments include, but are not limited to, such vectors as those containing the lac promoter (pUC-series); trp promoter (pBR322-trp); Ipp promoter (the pIN-series); lambda-pP or pR promoters (pOTS); or hybrid promoters such as ptac (pDR540). See Brosius, et al. (1988) "Expression Vectors Employing Lambda-, trp-, lac-,  
35 and Ipp-derived Promoters", in Vectors: A Survey of Molecular Cloning Vectors and Their Uses, (eds. Rodriguez and Denhardt), Butterworth, Boston, Chapter 10, pp. 205-236, which is incorporated herein by reference.

Lower eukaryotes, e.g., yeasts and Dictyostelium, may be transformed with DCRS8 sequence containing vectors. For purposes of this invention, the most common lower eukaryotic host is the baker's yeast, Saccharomyces cerevisiae. It will be used to generically represent lower eukaryotes although a number of other strains and species are also available. Yeast vectors typically consist of a replication origin (unless of the integrating type), a selection gene, a promoter, DNA encoding the receptor or its fragments, and sequences for translation termination, polyadenylation, and transcription termination. Suitable expression vectors for yeast include such constitutive promoters as 3-phosphoglycerate kinase and various other glycolytic enzyme gene promoters or such inducible promoters as the alcohol dehydrogenase 2 promoter or metallothionein promoter. Suitable vectors include derivatives of the following types: self-replicating low copy number (such as the YRp-series), self-replicating high copy number (such as the YE<sub>p</sub>-series); integrating types (such as the YIp-series), or mini-chromosomes (such as the YC<sub>p</sub>-series).

Higher eukaryotic tissue culture cells are normally the preferred host cells for expression of the functionally active interleukin or receptor proteins. In principle, many higher eukaryotic tissue culture cell lines are workable, e.g., insect baculovirus expression systems, whether from an invertebrate or vertebrate source. However, mammalian cells are preferred. Transformation or transfection and propagation of such cells has become a routine procedure. Examples of useful cell lines include HeLa cells, Chinese hamster ovary (CHO) cell lines, baby rat kidney (BRK) cell lines, insect cell lines, bird cell lines, and monkey (COS) cell lines. Expression vectors for such cell lines usually include an origin of replication, a promoter, a translation initiation site, RNA splice sites (if genomic DNA is used), a polyadenylation site, and a transcription termination site. These vectors also usually contain a selection gene or amplification gene. Suitable expression vectors may be plasmids, viruses, or retroviruses carrying promoters derived, e.g., from such sources as from adenovirus, SV40, parvoviruses, vaccinia virus, or cytomegalovirus. Representative examples of suitable expression vectors include pCDNA1; pCD, see Okayama, et al. (1985) Mol. Cell Biol. 5:1136-1142; pMC1neo PolyA, see Thomas, et al. (1987) Cell 51:503-512; and a baculovirus vector such as pAC 373 or pAC 610.

For secreted proteins and some membrane proteins, an open reading frame usually encodes a polypeptide that consists of a mature or secreted product covalently linked at its N-terminus to a signal peptide. The signal peptide is cleaved prior to secretion of the mature, or active, polypeptide. The cleavage site can be predicted with a high degree of accuracy from empirical rules, e.g., von-Heijne (1986) Nucleic Acids Research 14:4683-4690; and Nielsen, et al. (1997) Protein Eng. 10:1-12, and the precise amino acid composition of the signal peptide often does not appear to be critical to its function, e.g.,

Randall, et al. (1989) Science 243:1156-1159; and Kaiser, et al. (1987) Science 235:312-317. The mature proteins of the invention can be readily determined using standard methods.

5 It will often be desired to express these polypeptides in a system which provides a specific or defined glycosylation pattern. In this case, the usual pattern will be that provided naturally by the expression system. However, the pattern will be modifiable by exposing the polypeptide, e.g., an unglycosylated form, to appropriate glycosylating proteins introduced into a heterologous expression system. For example, the receptor gene may be co-transformed with one or more genes encoding mammalian or other  
10 glycosylating enzymes. Using this approach, certain mammalian glycosylation patterns will be achievable in prokaryote or other cells. Expression in prokaryote cells will typically lead to unglycosylated forms of protein.

The source of DCRS8 can be a eukaryotic or prokaryotic host expressing recombinant DCRS8, such as is described above. The source can also be a cell line, but  
15 other mammalian cell lines are also contemplated by this invention, with the preferred cell line being from the human species.

Now that the sequences are known, the primate DCRS8 or DCRS9, fragments, or derivatives thereof can be prepared by conventional processes for synthesizing peptides. These include processes such as are described in Stewart and Young (1984) Solid Phase Peptide Synthesis, Pierce Chemical Co., Rockford, IL; Bodanszky and Bodanszky (1984) The Practice of Peptide Synthesis, Springer-Verlag, New York; and Bodanszky (1984) The Principles of Peptide Synthesis, Springer-Verlag, New York; all of each which are incorporated herein by reference. For example, an azide process, an acid chloride process, an acid anhydride process, a mixed anhydride process, an active ester process  
25 (for example, p-nitrophenyl ester, N-hydroxysuccinimide ester, or cyanomethyl ester), a carbodiimidazole process, an oxidative-reductive process, or a dicyclohexylcarbodiimide (DCCD)/additive process can be used. Solid phase and solution phase syntheses are both applicable to the foregoing processes. Similar techniques can be used with partial DCRS8 or DCRS9 sequences.

30 The DCRS8 proteins, fragments, or derivatives are suitably prepared in accordance with the above processes as typically employed in peptide synthesis, generally either by a so-called stepwise process which comprises condensing an amino acid to the terminal amino acid, one by one in sequence, or by coupling peptide fragments to the terminal amino acid. Amino groups that are not being used in the coupling reaction  
35 typically must be protected to prevent coupling at an incorrect location.

If a solid phase synthesis is adopted, the C-terminal amino acid is bound to an insoluble carrier or support through its carboxyl group. The insoluble carrier is not

particularly limited as long as it has a binding capability to a reactive carboxyl group. Examples of such insoluble carriers include halomethyl resins, such as chloromethyl resin or bromomethyl resin, hydroxymethyl resins, phenol resins, tert-alkyloxycarbonylhydrazidated resins, and the like.

5           An amino group-protected amino acid is bound in sequence through condensation of its activated carboxyl group and the reactive amino group of the previously formed peptide or chain, to synthesize the peptide step by step. After synthesizing the complete sequence, the peptide is split off from the insoluble carrier to produce the peptide. This solid-phase approach is generally described by Merrifield, et al. (1963) in J. Am. Chem.  
10 Soc. 85:2149-2156, which is incorporated herein by reference.

          The prepared protein and fragments thereof can be isolated and purified from the reaction mixture by means of peptide separation, e.g., by extraction, precipitation, electrophoresis, various forms of chromatography, and the like. The receptors of this invention can be obtained in varying degrees of purity depending upon desired uses.  
15 Purification can be accomplished by use of the protein purification techniques disclosed herein, see below, or by the use of the antibodies herein described in methods of immunoabsorbant affinity chromatography. This immunoabsorbant affinity chromatography is carried out by first linking the antibodies to a solid support and then contacting the linked antibodies with solubilized lysates of appropriate cells, lysates of  
20 other cells expressing the receptor, or lysates or supernatants of cells producing the protein as a result of DNA techniques, see below.

          Generally, the purified protein will be at least about 40% pure, ordinarily at least about 50% pure, usually at least about 60% pure, typically at least about 70% pure, more typically at least about 80% pure, preferable at least about 90% pure and more preferably  
25 at least about 95% pure, and in particular embodiments, 97%-99% or more. Purity will usually be on a weight basis, but can also be on a molar basis. Different assays will be applied as appropriate. Individual proteins may be purified and thereafter combined.

## VI. Antibodies

30           Antibodies can be raised to the various mammalian, e.g., primate DCRS8 or DCRS9 proteins and fragments thereof, both in naturally occurring native forms and in their recombinant forms, the difference being that antibodies to the active receptor are more likely to recognize epitopes which are only present in the native conformations. Denatured antigen detection can also be useful in, e.g., Western analysis. Anti-idiotypic  
35 antibodies are also contemplated, which would be useful as agonists or antagonists of a natural receptor or an antibody.



Antibodies, including binding fragments and single chain versions, against predetermined fragments of the protein can be raised by immunization of animals with conjugates of the fragments with immunogenic proteins. Monoclonal antibodies are prepared from cells secreting the desired antibody. These antibodies can be screened for binding to normal or defective protein, or screened for agonistic or antagonistic activity. These monoclonal antibodies will usually bind with at least a  $K_D$  of about 1 mM, more usually at least about 300  $\mu$ M, typically at least about 100 $\mu$ M, more typically at least about 30  $\mu$ M, preferably at least about 10  $\mu$ M, and more preferably at least about 3  $\mu$ M or better.

The antibodies, including antigen binding fragments, of this invention can have significant diagnostic or therapeutic value. They can be potent antagonists that bind to the receptor and inhibit binding to ligand or inhibit the ability of the receptor to elicit a biological response, e.g., act on its substrate. They also can be useful as non-neutralizing antibodies and can be coupled to toxins or radionuclides to bind producing cells, or cells localized to the source of the interleukin. Further, these antibodies can be conjugated to drugs or other therapeutic agents, either directly or indirectly by means of a linker.

The antibodies of this invention can also be useful in diagnostic applications. As capture or non-neutralizing antibodies, they might bind to the receptor without inhibiting ligand or substrate binding. As neutralizing antibodies, they can be useful in competitive binding assays. They will also be useful in detecting or quantifying ligand. They may be used as reagents for Western blot analysis, or for immunoprecipitation or immunopurification of the respective protein. Likewise, nucleic acids and proteins may be immobilized to solid substrates for affinity purification or detection methods. The substrates may be, e.g., solid resin beads or sheets of plastic.

Protein fragments may be joined to other materials, particularly polypeptides, as fused or covalently joined polypeptides to be used as immunogens. Mammalian cytokine receptors and fragments may be fused or covalently linked to a variety of immunogens, such as keyhole limpet hemocyanin, bovine serum albumin, tetanus toxoid, etc. See (1969) Microbiology, Hoeber Medical Division, Harper and Row; Landsteiner (1962) Specificity of Serological Reactions, Dover Publications, New York; and Williams, et al. (1967) Methods in Immunology and Immunochemistry, Vol. 1, Academic Press, New York; each of which is incorporated herein by reference, for descriptions of methods of preparing polyclonal antisera. A typical method involves hyperimmunization of an animal with an antigen. The blood of the animal is then collected shortly after the repeated immunizations and the gamma globulin is isolated.

In some instances, it is desirable to prepare monoclonal antibodies from various mammalian hosts, such as mice, rodents, primates, humans, etc. Description of

techniques for preparing such monoclonal antibodies may be found in, e.g., Stites, et al. (eds.) Basic and Clinical Immunology (4th ed.), Lange Medical Publications, Los Altos, CA, and references cited therein; Harlow and Lane (1988) Antibodies: A Laboratory Manual, CSH Press; Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed.) Academic Press, New York; and particularly in Kohler and Milstein (1975) Nature 256:495-497, which discusses one method of generating monoclonal antibodies. Each of these references is incorporated herein by reference. Summarized briefly, this method involves injecting an animal with an immunogen. The animal is then sacrificed and cells taken from its spleen, which are then fused with myeloma cells. The result is a hybrid cell or "hybridoma" that is capable of reproducing in vitro. The population of hybridomas is then screened to isolate individual clones, each of which secrete a single antibody species to the immunogen. In this manner, the individual antibody species obtained are the products of immortalized and cloned single B cells from the immune animal generated in response to a specific site recognized on the immunogenic substance.

Other suitable techniques involve in vitro exposure of lymphocytes to the antigenic polypeptides or alternatively to selection of libraries of antibodies in phage or similar vectors. See, Huse, et al. (1989) "Generation of a Large Combinatorial Library of the Immunoglobulin Repertoire in Phage Lambda," Science 246:1275-1281; and Ward, et al. (1989) Nature 341:544-546, each of which is incorporated herein by reference. The polypeptides and antibodies of the present invention may be used with or without modification, including chimeric or humanized antibodies. Frequently, the polypeptides and antibodies will be labeled by joining, either covalently or non-covalently, a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent moieties, chemiluminescent moieties, magnetic particles, and the like. Patents, teaching the use of such labels include U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241. Also, recombinant or chimeric immunoglobulins may be produced, see Cabilly, U.S. Patent No. 4,816,567; or made in transgenic mice, see Mendez, et al. (1997) Nature Genetics 15:146-156; Abgenix; and Medarex. These references are incorporated herein by reference.

The antibodies of this invention can also be used for affinity chromatography in isolating the DCRS8 proteins or peptides. Columns can be prepared where the antibodies are linked to a solid support, e.g., particles, such as agarose, Sephadex, or the like, where a cell lysate may be passed through the column, the column washed, followed by increasing concentrations of a mild denaturant, whereby the purified protein will be

released. Alternatively, the protein may be used to purify antibody. Appropriate cross absorptions or depletions may be applied.

5 The antibodies may also be used to screen expression libraries for particular expression products. Usually the antibodies used in such a procedure will be labeled with a moiety allowing easy detection of presence of antigen by antibody binding.

Antibodies raised against a cytokine receptor will also be used to raise anti-idiotypic antibodies. These will be useful in detecting or diagnosing various immunological conditions related to expression of the protein or cells which express the protein. They also will be useful as agonists or antagonists of the ligand, which may be competitive inhibitors or substitutes for naturally occurring ligands.

10 A cytokine receptor protein that specifically binds to or that is specifically immunoreactive with an antibody generated against a defined immunogen, such as an immunogen consisting of the amino acid sequence of SEQ ID NO: 14, is typically determined in an immunoassay. The immunoassay typically uses a polyclonal antiserum which was raised, e.g., to a protein of SEQ ID NO: 14. This antiserum is selected to have low crossreactivity against other cytokine receptor family members, preferably from the same species, and any such crossreactivity is removed by immunoabsorption prior to use in the immunoassay.

15 In order to produce antisera for use in an immunoassay, the protein, e.g., of SEQ ID NO: 14, is isolated as described herein. For example, recombinant protein may be produced in a mammalian cell line. An appropriate host, e.g., an inbred strain of mice such as Balb/c, is immunized with the selected protein, typically using a standard adjuvant, such as Freund's adjuvant, and a standard mouse immunization protocol (see Harlow and Lane, supra). Alternatively, a synthetic peptide derived from the sequences disclosed herein and conjugated to a carrier protein can be used as an immunogen.

25 Polyclonal sera are collected and titered against the immunogen protein in an immunoassay, e.g., a solid phase immunoassay with the immunogen immobilized on a solid support. Polyclonal antisera with a titer of  $10^4$  or greater are selected and tested for their cross reactivity against other cytokine receptor family members using a competitive binding immunoassay such as the one described in Harlow and Lane, supra, at pages 570-573. Preferably at least two cytokine receptor family members are used in this determination. These cytokine receptor family members can be produced as recombinant proteins and isolated using standard molecular biology and protein chemistry techniques as described herein.

30 Immunoassays in the competitive binding format can be used for the crossreactivity determinations. For example, the protein of SEQ ID NO: 14 can be immobilized to a solid support. Proteins added to the assay compete with the binding of

the antisera to the immobilized antigen. The ability of the above proteins to compete with the binding of the antisera to the immobilized protein is compared to the other proteins. The percent crossreactivity for the above proteins is calculated, using standard calculations. Those antisera with less than 10% crossreactivity with each of the proteins listed above are selected and pooled. The cross-reacting antibodies are then removed from the pooled antisera by immunoabsorption with the above-listed proteins.

The immunoabsorbed and pooled antisera are then used in a competitive binding immunoassay as described above to compare a second protein to the immunogen protein (e.g., the DCRS8 like protein of SEQ ID NO: 14). In order to make this comparison, the two proteins are each assayed at a wide range of concentrations and the amount of each protein required to inhibit 50% of the binding of the antisera to the immobilized protein is determined. If the amount of the second protein required is less than twice the amount of the protein of the selected protein or proteins that is required, then the second protein is said to specifically bind to an antibody generated to the immunogen.

It is understood that these cytokine receptor proteins are members of a family of homologous proteins that comprise at least 9 so far identified members, 6 mammalian and 3 worm embodiments. For a particular gene product, such as the DCRS8, the term refers not only to the amino acid sequences disclosed herein, but also to other proteins that are allelic, non-allelic, or species variants. It is also understood that the terms include nonnatural mutations introduced by deliberate mutation using conventional recombinant technology such as single site mutation, or by excising short sections of DNA encoding the respective proteins, or by substituting new amino acids, or adding new amino acids. Such minor alterations typically will substantially maintain the immunoidentity of the original molecule and/or its biological activity. Thus, these alterations include proteins that are specifically immunoreactive with a designated naturally occurring DCRS8 protein. The biological properties of the altered proteins can be determined by expressing the protein in an appropriate cell line and measuring the appropriate effect, e.g., upon transfected lymphocytes. Particular protein modifications considered minor would include conservative substitution of amino acids with similar chemical properties, as described above for the cytokine receptor family as a whole. By aligning a protein optimally with the protein of the cytokine receptors and by using the conventional immunoassays described herein to determine immunoidentity, one can determine the protein compositions of the invention.

## VII. Kits and quantitation

Both naturally occurring and recombinant forms of the cytokine receptor like molecules of this invention are particularly useful in kits and assay methods. For

example, these methods would also be applied to screening for binding activity, e.g., ligands for these proteins. Several methods of automating assays have been developed in recent years so as to permit screening of tens of thousands of compounds per year. See, e.g., a BIOMEK automated workstation, Beckman Instruments, Palo Alto, California, and Fodor, et al. (1991) Science 251:767-773, which is incorporated herein by reference. The latter describes means for testing binding by a plurality of defined polymers synthesized on a solid substrate. The development of suitable assays to screen for a ligand or agonist/antagonist homologous proteins can be greatly facilitated by the availability of large amounts of purified, soluble cytokine receptors in an active state such as is provided by this invention.

Purified protein can be coated directly onto plates for use in the aforementioned ligand screening techniques. However, non-neutralizing antibodies to these proteins can be used as capture antibodies to immobilize the respective receptor on the solid phase, useful, e.g., in diagnostic uses.

This invention also contemplates use of receptor subunit, fragments thereof, peptides, and their fusion products in a variety of diagnostic kits and methods for detecting the presence of the protein or its ligand. Alternatively, or additionally, antibodies against the molecules may be incorporated into the kits and methods. Typically the kit will have a compartment containing, e.g., a DCRS8 peptide or gene segment or a reagent which recognizes one or the other. Typically, recognition reagents, in the case of peptide, would be a receptor or antibody, or in the case of a gene segment, would usually be a hybridization probe.

A preferred kit for determining the concentration of DCRS8 in a sample would typically comprise a labeled compound, e.g., ligand or antibody, having known binding affinity for DCRS8, a source of DCRS8 (naturally occurring or recombinant) as a positive control, and a means for separating the bound from free labeled compound, e.g., a solid phase for immobilizing the DCRS8 in the test sample. Compartments containing reagents, and instructions, will normally be provided. Appropriate nucleic acid or protein containing kits are also provided.

Antibodies, including antigen binding fragments, specific for mammalian DCRS8 or a peptide fragment, or receptor fragments are useful in diagnostic applications to detect the presence of elevated levels of ligand and/or its fragments. Diagnostic assays may be homogeneous (without a separation step between free reagent and antibody-antigen complex) or heterogeneous (with a separation step). Various commercial assays exist, such as radioimmunoassay (RIA), enzyme-linked immunosorbent assay (ELISA), enzyme immunoassay (EIA), enzyme-multiplied immunoassay technique (EMIT), substrate-labeled fluorescent immunoassay (SLFIA) and the like. For example, unlabeled

antibodies can be employed by using a second antibody which is labeled and which recognizes the antibody to a cytokine receptor or to a particular fragment thereof. These assays have also been extensively discussed in the literature. See, e.g., Harlow and Lane (1988) Antibodies: A Laboratory Manual, CSH, and Coligan (ed. 1991 and periodic supplements) Current Protocols In Immunology Greene/Wiley, New York.

Anti-idiotypic antibodies may have similar use to serve as agonists or antagonists of cytokine receptors. These should be useful as therapeutic reagents under appropriate circumstances.

Frequently, the reagents for diagnostic assays are supplied in kits, so as to optimize the sensitivity of the assay. For the subject invention, depending upon the nature of the assay, the protocol, and the label, either labeled or unlabeled antibody, or labeled ligand is provided. This is usually in conjunction with other additives, such as buffers, stabilizers, materials necessary for signal production such as substrates for enzymes, and the like. Preferably, the kit will also contain instructions for proper use and disposal of the contents after use. Typically the kit has compartments for each useful reagent, and will contain instructions for proper use and disposal of reagents. Desirably, the reagents are provided as a dry lyophilized powder, where the reagents may be reconstituted in an aqueous medium having appropriate concentrations for performing the assay.

The aforementioned constituents of the diagnostic assays may be used without modification or may be modified in a variety of ways. For example, labeling may be achieved by covalently or non-covalently joining a moiety which directly or indirectly provides a detectable signal. In many of these assays, a test compound, cytokine receptor, or antibodies thereto can be labeled either directly or indirectly. Possibilities for direct labeling include label groups: radiolabels such as  $^{125}\text{I}$ , enzymes (U.S. Pat. No. 3,645,090) such as peroxidase and alkaline phosphatase, and fluorescent labels (U.S. Pat. No. 3,940,475) capable of monitoring the change in fluorescence intensity, wavelength shift, or fluorescence polarization. Both of the patents are incorporated herein by reference. Possibilities for indirect labeling include biotinylation of one constituent followed by binding to avidin coupled to one of the above label groups.

There are also numerous methods of separating the bound from the free ligand, or alternatively the bound from the free test compound. The cytokine receptor can be immobilized on various matrixes followed by washing. Suitable matrices include plastic such as an ELISA plate, filters, and beads. Methods of immobilizing the receptor to a matrix include, without limitation, direct adhesion to plastic, use of a capture antibody, chemical coupling, and biotin-avidin. The last step in this approach involves the precipitation of antibody/antigen complex by any of several methods including those

utilizing, e.g., an organic solvent such as polyethylene glycol or a salt such as ammonium sulfate. Other suitable separation techniques include, without limitation, the fluorescein antibody magnetizable particle method described in Rattle, et al. (1984) Clin. Chem. 30(9):1457-1461, and the double antibody magnetic particle separation as described in  
5 U.S. Pat. No. 4,659,678, each of which is incorporated herein by reference.

The methods for linking protein or fragments to various labels have been extensively reported in the literature and do not require detailed discussion here. Many of the techniques involve the use of activated carboxyl groups either through the use of carbodiimide or active esters to form peptide bonds, the formation of thioethers by  
10 reaction of a mercapto group with an activated halogen such as chloroacetyl, or an activated olefin such as maleimide, for linkage, or the like. Fusion proteins will also find use in these applications.

Another diagnostic aspect of this invention involves use of oligonucleotide or polynucleotide sequences taken from the sequence of an cytokine receptor. These  
15 sequences can be used as probes for detecting levels of the respective cytokine receptor in patients suspected of having an immunological disorder. The preparation of both RNA and DNA nucleotide sequences, the labeling of the sequences, and the preferred size of the sequences has received ample description and discussion in the literature. Normally an oligonucleotide probe should have at least about 14 nucleotides, usually at least about  
20 18 nucleotides, and the polynucleotide probes may be up to several kilobases. Various labels may be employed, most commonly radionuclides, particularly  $^{32}\text{P}$ . However, other techniques may also be employed, such as using biotin modified nucleotides for introduction into a polynucleotide. The biotin then serves as the site for binding to avidin or antibodies, which may be labeled with a wide variety of labels, such as radionuclides,  
25 fluorescers, enzymes, or the like. Alternatively, antibodies may be employed which can recognize specific duplexes, including DNA duplexes, RNA duplexes, DNA-RNA hybrid duplexes, or DNA-protein duplexes. The antibodies in turn may be labeled and the assay carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected. The use of  
30 probes to the novel RNA may be carried out in conventional techniques such as nucleic acid hybridization, plus and minus screening, recombinational probing, hybrid released translation (HRT), and hybrid arrested translation (HART). Antisense nucleic acids, which may be used to block protein expression, are also provided. See, e.g., Isis Pharmaceuticals, Sequitur, Inc., or Hybridon. This also includes amplification techniques  
35 such as polymerase chain reaction (PCR).

Diagnostic kits which also test for the qualitative or quantitative presence of other markers are also contemplated. Diagnosis or prognosis may depend on the combination

of multiple indications used as markers. Thus, kits may test for combinations of markers. See, e.g., Viallet, et al. (1989) Progress in Growth Factor Res. 1:89-97.

#### VIII. Therapeutic Utility

5           This invention provides reagents with significant therapeutic value. See, e.g., Levitzki (1996) Curr. Opin. Cell Biol. 8:239-244. The cytokine receptors (naturally occurring or recombinant), fragments thereof, mutein receptors, and antibodies, along with compounds identified as having binding affinity to the receptors or antibodies, should be useful in the treatment of conditions exhibiting abnormal expression of the  
10       receptors of their ligands. Such abnormality will typically be manifested by immunological disorders, e.g., innate immunity, or developmentally. Additionally, this invention should provide therapeutic value in various diseases or disorders associated with abnormal expression or abnormal triggering of response to the ligand. For example, the IL-1 ligands have been suggested to be involved in morphologic development, e.g.,  
15       dorso-ventral polarity determination, and immune responses, particularly the primitive innate responses. See, e.g., Sun, et al. (1991) Eur. J. Biochem. 196:247-254; and Hultmark (1994) Nature 367:116-117.

          Recombinant cytokine receptors, muteins, agonist or antagonist antibodies thereto, or antibodies can be purified and then administered to a patient. These reagents can be  
20       combined for therapeutic use with additional active ingredients, e.g., in conventional pharmaceutically acceptable carriers or diluents, along with physiologically innocuous stabilizers and excipients. These combinations can be sterile, e.g., filtered, and placed into dosage forms as by lyophilization in dosage vials or storage in stabilized aqueous preparations. This invention also contemplates use of antibodies or binding fragments  
25       thereof which are not complement binding.

          Ligand screening using cytokine receptor or fragments thereof can be performed to identify molecules having binding affinity to the receptors. Subsequent biological assays can then be utilized to determine if a putative ligand can provide competitive binding, which can block intrinsic stimulating activity. Receptor fragments can be used  
30       as a blocker or antagonist in that it blocks the activity of ligand. Likewise, a compound having intrinsic stimulating activity can activate the receptor and is thus an agonist in that it simulates the activity of ligand, e.g., inducing signaling. This invention further contemplates the therapeutic use of antibodies to cytokine receptors as antagonists.

          The quantities of reagents necessary for effective therapy will depend upon many  
35       different factors, including means of administration, target site, reagent physiological life, pharmacological life, physiological state of the patient, and other medicants administered. Thus, treatment dosages should be titrated to optimize safety and efficacy. Typically,



dosages used in vitro may provide useful guidance in the amounts useful for in situ administration of these reagents. Animal testing of effective doses for treatment of particular disorders will provide further predictive indication of human dosage. Various considerations are described, e.g., in Gilman, et al. (eds. 1990) Goodman and Gilman's: The Pharmacological Bases of Therapeutics, 8th Ed., Pergamon Press; and Remington's Pharmaceutical Sciences, 17th ed. (1990), Mack Publishing Co., Easton, Penn.; each of which is hereby incorporated herein by reference. Methods for administration are discussed therein and below, e.g., for oral, intravenous, intraperitoneal, or intramuscular administration, transdermal diffusion, and others. Pharmaceutically acceptable carriers will include water, saline, buffers, and other compounds described, e.g., in the Merck Index, Merck & Co., Rahway, New Jersey. Because of the likely high affinity binding, or turnover numbers, between a putative ligand and its receptors, low dosages of these reagents would be initially expected to be effective. And the signaling pathway suggests extremely low amounts of ligand may have effect. Thus, dosage ranges would ordinarily be expected to be in amounts lower than 1 mM concentrations, typically less than about 10  $\mu$ M concentrations, usually less than about 100 nM, preferably less than about 10 pM (picomolar), and most preferably less than about 1 fM (femtomolar), with an appropriate carrier. Slow release formulations, or slow release apparatus will often be utilized for continuous administration.

Cytokine receptors, fragments thereof, and antibodies or its fragments, antagonists, and agonists, may be administered directly to the host to be treated or, depending on the size of the compounds, it may be desirable to conjugate them to carrier proteins such as ovalbumin or serum albumin prior to their administration. Therapeutic formulations may be administered in many conventional dosage formulations. While it is possible for the active ingredient to be administered alone, it is preferable to present it as a pharmaceutical formulation. Formulations comprise at least one active ingredient, as defined above, together with one or more acceptable carriers thereof. Each carrier must be both pharmaceutically and physiologically acceptable in the sense of being compatible with the other ingredients and not injurious to the patient. Formulations include those suitable for oral, rectal, nasal, or parenteral (including subcutaneous, intramuscular, intravenous and intradermal) administration. The formulations may conveniently be presented in unit dosage form and may be prepared by methods well known in the art of pharmacy. See, e.g., Gilman, et al. (eds. 1990) Goodman and Gilman's: The Pharmacological Bases of Therapeutics, 8th Ed., Pergamon Press; and Remington's Pharmaceutical Sciences, 17th ed. (1990), Mack Publishing Co., Easton, Penn.; Avis, et al. (eds. 1993) Pharmaceutical Dosage Forms: Parenteral Medications Dekker, NY; Lieberman, et al. (eds. 1990) Pharmaceutical Dosage Forms: Tablets Dekker, NY; and

Lieberman, et al. (eds. 1990) Pharmaceutical Dosage Forms: Disperse Systems Dekker, NY. The therapy of this invention may be combined with or used in association with other therapeutic agents, particularly agonists or antagonists of other cytokine receptor family members.

5

## IX. Screening

Drug screening using DCRS8 or fragments thereof can be performed to identify compounds having binding affinity to the receptor subunit, including isolation of associated components. Subsequent biological assays can then be utilized to determine if the compound has intrinsic stimulating activity and is therefore a blocker or antagonist in that it blocks the activity of the ligand. Likewise, a compound having intrinsic stimulating activity can activate the receptor and is thus an agonist in that it simulates the activity of a cytokine ligand. This invention further contemplates the therapeutic use of antibodies to the receptor as cytokine agonists or antagonists.

10

Similarly, complexes comprising multiple proteins may be used to screen for ligands or reagents capable of recognizing the complex. Most cytokine receptors comprise at least two subunits, which may be the same, or distinct. Alternatively, the transmembrane receptor may bind to a complex comprising a cytokine-like ligand associated with another soluble protein serving, e.g., as a second receptor subunit.

15

One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant DNA molecules expressing the DCRS8 in combination with another cytokine receptor subunit. Cells may be isolated which express a receptor in isolation from other functional receptors. Such cells, either in viable or fixed form, can be used for standard antibody/antigen or ligand/receptor binding assays. See also, Parce, et al. (1989) Science 246:243-247; and Owicki, et al. (1990) Proc. Nat'l Acad. Sci. USA 87:4007-4011, which describe sensitive methods to detect cellular responses. Competitive assays are particularly useful, where the cells (source of putative ligand) are contacted and incubated with a labeled receptor or antibody having known binding affinity to the ligand, such as  $^{125}\text{I}$ -antibody, and a test sample whose binding affinity to the binding composition is being measured. The bound and free labeled binding compositions are then separated to assess the degree of ligand binding. The amount of test compound bound is inversely proportional to the amount of labeled receptor binding to the known source. Many techniques can be used to separate bound from free ligand to assess the degree of ligand binding. This separation step could typically involve a procedure such as adhesion to filters followed by washing, adhesion to plastic followed by washing, or centrifugation of the cell membranes. Viable cells could also be used to screen for the effects of drugs on cytokine mediated functions, e.g., second messenger

20

25

30

35

levels, e.g.,  $\text{Ca}^{++}$ ; cell proliferation; inositol phosphate pool changes; and others. Some detection methods allow for elimination of a separation step, e.g., a proximity sensitive detection system. Calcium sensitive dyes will be useful for detecting  $\text{Ca}^{++}$  levels, with a fluorimeter or a fluorescence cell sorting apparatus.

5

#### X. Ligands

The descriptions of the DCRS8 herein provides means to identify ligands, as described above. Such ligand should bind specifically to the respective receptor with reasonably high affinity. Various constructs are made available which allow either labeling of the receptor to detect its ligand. For example, directly labeling cytokine receptor, fusing onto it markers for secondary labeling, e.g., FLAG or other epitope tags, etc., will allow detection of receptor. This can be histological, as an affinity method for biochemical purification, or labeling or selection in an expression cloning approach. A two-hybrid selection system may also be applied making appropriate constructs with the available cytokine receptor sequences. See, e.g., Fields and Song (1989) Nature 340:245-246.

10

15

Most likely candidates will be structurally related to members of the IL-17 family. See, e.g., USSN 09/480,287.

The broad scope of this invention is best understood with reference to the following examples, which are not intended to limit the inventions to the specific embodiments.

20

### EXAMPLES

#### I. General Methods

25

Some of the standard methods are described or referenced, e.g., in Maniatis, et al. (1982) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor Press; Sambrook, et al. (1989) Molecular Cloning: A Laboratory Manual, (2d ed.), vols. 1-3, CSH Press, NY; or Ausubel, et al. (1987 and Supplements) Current Protocols in Molecular Biology, Greene/Wiley, New York. Methods for protein purification include such methods as ammonium sulfate precipitation, column chromatography, electrophoresis, centrifugation, crystallization, and others. See, e.g., Ausubel, et al. (1987 and periodic supplements); Coligan, et al. (ed. 1996) and periodic supplements, Current Protocols In Protein Science Greene/Wiley, New York; Deutscher (1990) "Guide to Protein Purification" in Methods in Enzymology, vol. 182, and other volumes in this series; and manufacturer's literature on use of protein purification products, e.g., Pharmacia, Piscataway, N.J., or Bio-Rad, Richmond, CA. Combination

30

35

with recombinant techniques allow fusion to appropriate segments, e.g., to a FLAG sequence or an equivalent which can be fused via a protease-removable sequence. See, e.g., Hochuli (1990) "Purification of Recombinant Proteins with Metal Chelate Absorbent" in Setlow (ed.) Genetic Engineering, Principle and Methods 12:87-98, Plenum Press, N.Y.; and Crowe, et al. (1992) QIAexpress: The High Level Expression & Protein Purification System QUIAGEN, Inc., Chatsworth, CA.

Computer sequence analysis is performed, e.g., using available software programs, including those from the GCG (U. Wisconsin) and GenBank sources. Public sequence databases were also used, e.g., from GenBank and others.

Many techniques applicable to IL-10 receptors may be applied to the DCRSs, as described, e.g., in USSN 08/110,683 (IL-10 receptor), which is incorporated herein by reference.

## II. Computational Analysis

Human sequences related to cytokine receptors were identified from genomic sequence database using, e.g., the BLAST server (Altschul, et al. (1994) Nature Genet. 6:119-129). Standard analysis programs may be used to evaluate structure, e.g., PHD (Rost and Sander (1994) Proteins 19:55-72) and DSC (King and Sternberg (1996) Protein Sci. 5:2298-2310). Standard comparison software includes, e.g., Altschul, et al. (1990) J. Mol. Biol. 215:403-10; Waterman (1995) Introduction to Computational Biology: Maps, Sequences, and Genomes Chapman & Hall; Lander and Waterman (eds. 1995) Calculating the Secrets of Life: Applications of the Mathematical Sciences in Molecular Biology National Academy Press; and Speed and Waterman (eds. 1996) Genetic Mapping and DNA Sequencing (IMA Volumes in Mathematics and Its Applications, Vol 81) Springer Verlag. Each reference is incorporate herein by reference.

## III. Cloning of full-length cDNAs; Chromosomal localization

PCR primers derived from the sequences are used to probe a human cDNA library. Sequences may be derived, e.g., from Tables 1-5, preferably those adjacent the ends of sequences. Full length cDNAs for primate, rodent, or other species DCRS8 are cloned, e.g., by DNA hybridization screening of  $\lambda$ gt10 phage. PCR reactions are conducted using T. aquaticus Taqplus DNA polymerase (Stratagene) under appropriate conditions. Extending partial length cDNA clones is typically routine.

Chromosome spreads are prepared. In situ hybridization is performed on chromosome preparations obtained from phytohemagglutinin-stimulated human lymphocytes cultured for 72 h. 5-bromodeoxyuridine was added for the final seven hours

of culture (60 µg/ml of medium), to ensure a posthybridization chromosomal banding of good quality.

A PCR fragment, amplified with the help of primers, is cloned into an appropriate vector. The vector is labeled by nick-translation with  $^3\text{H}$ . The radiolabeled probe is  
5 hybridized to metaphase spreads at final concentration of 200 ng/ml of hybridization solution as described, e.g., in Mattei, et al. (1985) Hum. Genet. 69:327-331.

After coating with nuclear track emulsion (KODAK NTB<sub>2</sub>), slides are exposed. To avoid any slipping of silver grains during the banding procedure, chromosome spreads are first stained with buffered Giemsa solution and metaphase photographed. R-banding  
10 is then performed by the fluorochrome-photolysis-Giemsa (FPG) method and metaphases rephotographed before analysis.

Similar appropriate methods are used for other species.

#### IV. Localization of mRNA

Human multiple tissue (Cat# 1, 2) and cancer cell line blots (Cat# 7757-1),  
15 containing approximately 2 µg of poly(A)<sup>+</sup> RNA per lane, are purchased from Clontech (Palo Alto, CA). Probes are radiolabeled with [ $\alpha$ - $^{32}\text{P}$ ] dATP, e.g., using the Amersham Rediprime random primer labeling kit (RPN1633). Prehybridization and hybridizations are performed, e.g., at 65° C in 0.5 M Na<sub>2</sub>HPO<sub>4</sub>, 7% SDS, 0.5 M EDTA (pH 8.0). High  
20 stringency washes are conducted, e.g., at 65° C with two initial washes in 2 x SSC, 0.1% SDS for 40 min followed by a subsequent wash in 0.1 x SSC, 0.1% SDS for 20 min. Membranes are then exposed at -70° C to X-Ray film (Kodak) in the presence of intensifying screens. More detailed studies by cDNA library Southern blots are performed with selected appropriate human DCRS clones to examine their expression in  
25 hemopoietic or other cell subsets.

Alternatively, two appropriate primers are selected from Tables 1-5. RT-PCR is used on an appropriate mRNA sample selected for the presence of message to produce a cDNA, e.g., a sample which expresses the gene.

Full length clones may be isolated by hybridization of cDNA libraries from  
30 appropriate tissues pre-selected by PCR signal. Northern blots can be performed.

Message for genes encoding DCRS will be assayed by appropriate technology, e.g., PCR, immunoassay, hybridization, or otherwise. Tissue and organ cDNA preparations are available, e.g., from Clontech, Mountain View, CA. Identification of sources of natural expression are useful, as described. And the identification of functional  
35 receptor subunit pairings will allow for prediction of what cells express the combination of receptor subunits which will result in a physiological responsiveness to each of the cytokine ligands.

For mouse counterpart distribution, e.g., Southern Analysis can be performed: DNA (5 µg) from a primary amplified cDNA library was digested with appropriate restriction enzymes to release the inserts, run on a 1% agarose gel and transferred to a nylon membrane (Schleicher and Schuell, Keene, NH).

- 5 Samples for mouse mRNA isolation may include: resting mouse fibroblastic L cell line (C200); Braf:ER (Braf fusion to estrogen receptor) transfected cells, control (C201); T cells, TH1 polarized (Mel14 bright, CD4+ cells from spleen, polarized for 7 days with IFN-γ and anti IL-4; T200); T cells, TH2 polarized (Mel14 bright, CD4+ cells from spleen, polarized for 7 days with IL-4 and anti-IFN-γ; T201); T cells, highly TH1 polarized (see Openshaw, et al. (1995) J. Exp. Med. 182:1357-1367; activated with anti-CD3 for 2, 6, 16 h pooled; T202); T cells, highly TH2 polarized (see Openshaw, et al. (1995) J. Exp. Med. 182:1357-1367; activated with anti-CD3 for 2, 6, 16 h pooled; T203); CD44- CD25+ pre T cells, sorted from thymus (T204); TH1 T cell clone D1.1, resting for 3 weeks after last stimulation with antigen (T205); TH1 T cell clone D1.1, 10  
10 µg/ml ConA stimulated 15 h (T206); TH2 T cell clone CDC35, resting for 3 weeks after last stimulation with antigen (T207); TH2 T cell clone CDC35, 10 µg/ml ConA stimulated 15 h (T208); Mel14+ naive T cells from spleen, resting (T209); Mel14+ T cells, polarized to Th1 with IFN-γ/IL-12/anti-IL-4 for 6, 12, 24 h pooled (T210); Mel14+ T cells, polarized to Th2 with IL-4/anti-IFN-γ for 6, 13, 24 h pooled (T211); unstimulated  
15 mature B cell leukemia cell line A20 (B200); unstimulated B cell line CH12 (B201); unstimulated large B cells from spleen (B202); B cells from total spleen, LPS activated (B203); metrizamide enriched dendritic cells from spleen, resting (D200); dendritic cells from bone marrow, resting (D201); monocyte cell line RAW 264.7 activated with LPS 4 h (M200); bone-marrow macrophages derived with GM and M-CSF (M201); macrophage  
20 cell line J774, resting (M202); macrophage cell line J774 + LPS + anti-IL-10 at 0.5, 1, 3, 6, 12 h pooled (M203); macrophage cell line J774 + LPS + IL-10 at 0.5, 1, 3, 5, 12 h pooled (M204); aerosol challenged mouse lung tissue, Th2 primers, aerosol OVA challenge 7, 14, 23 h pooled (see Garlisi, et al. (1995) Clinical Immunology and Immunopathology 75:75-83; X206); Nippostrongylus-infected lung tissue (see Coffman, et al. (1989) Science 245:308-310; X200); total adult lung, normal (O200); total lung, rag-1 (see Schwarz, et al. (1993) Immunodeficiency 4:249-252; O205); IL-10 K.O. spleen (see Kuhn, et al. (1991) Cell 75:263-274; X201); total adult spleen, normal (O201); total spleen, rag-1 (O207); IL-10 K.O. Peyer's patches (O202); total Peyer's patches, normal (O210); IL-10 K.O. mesenteric lymph nodes (X203); total mesenteric lymph nodes,  
25 normal (O211); IL-10 K.O. colon (X203); total colon, normal (O212); NOD mouse pancreas (see Makino, et al. (1980) Jikken Dobutsu 29:1-13; X205); total thymus, rag-1 (O208); total kidney, rag-1 (O209); total heart, rag-1 (O202); total brain, rag-1 (O203);  
30

total testes, rag-1 (O204); total liver, rag-1 (O206); rat normal joint tissue (O300); and rat arthritic joint tissue (X300).

Samples for human mRNA isolation may include, e.g.: peripheral blood mononuclear cells (monocytes, T cells, NK cells, granulocytes, B cells), resting (T100);  
 5 peripheral blood mononuclear cells, activated with anti-CD3 for 2, 6, 12 h pooled (T101); T cell, TH0 clone Mot 72, resting (T102); T cell, TH0 clone Mot 72, activated with anti-CD28 and anti-CD3 for 3, 6, 12 h pooled (T103); T cell, TH0 clone Mot 72, anergic treated with specific peptide for 2, 7, 12 h pooled (T104); T cell, TH1 clone HY06, resting (T107); T cell, TH1 clone HY06, activated with anti-CD28 and anti-CD3 for 3, 6,  
 10 12 h pooled (T108); T cell, TH1 clone HY06, anergic treated with specific peptide for 2, 6, 12 h pooled (T109); T cell, TH2 clone HY935, resting (T110); T cell, TH2 clone HY935, activated with anti-CD28 and anti-CD3 for 2, 7, 12 h pooled (T111); T cells CD4+CD45RO- T cells polarized 27 days in anti-CD28, IL-4, and anti IFN- $\gamma$ , TH2 polarized, activated with anti-CD3 and anti-CD28 4 h (T116); T cell tumor lines Jurkat and Hut78, resting (T117); T cell clones, pooled AD130.2, Tc783.12, Tc783.13,  
 15 Tc783.58, Tc782.69, resting (T118); T cell random  $\gamma\delta$  T cell clones, resting (T119); Splenocytes, resting (B100); Splenocytes, activated with anti-CD40 and IL-4 (B101); B cell EBV lines pooled WT49, RSB, JY, CVIR, 721.221, RM3, HSY, resting (B102); B cell line JY, activated with PMA and ionomycin for 1, 6 h pooled (B103); NK 20 clones pooled, resting (K100); NK 20 clones pooled, activated with PMA and ionomycin for 6 h (K101); NKL clone, derived from peripheral blood of LGL leukemia patient, IL-2 treated (K106); NK cytotoxic clone 640-A30-1, resting (K107); hematopoietic precursor line TF1, activated with PMA and ionomycin for 1, 6 h pooled (C100); U937 premonocytic line, resting (M100); U937 premonocytic line, activated with PMA and ionomycin for 1,  
 20 6 h pooled (M101); elutriated monocytes, activated with LPS, IFN $\gamma$ , anti-IL-10 for 1, 2, 6, 12, 24 h pooled (M102); elutriated monocytes, activated with LPS, IFN $\gamma$ , IL-10 for 1, 2, 6, 12, 24 h pooled (M103); elutriated monocytes, activated with LPS, IFN $\gamma$ , anti-IL-10 for 4, 16 h pooled (M106); elutriated monocytes, activated with LPS, IFN $\gamma$ , IL-10 for 4, 16 h pooled (M107); elutriated monocytes, activated LPS for 1 h (M108); elutriated  
 30 monocytes, activated LPS for 6 h (M109); DC 70% CD1a+, from CD34+ GM-CSF, TNF $\alpha$  12 days, resting (D101); DC 70% CD1a+, from CD34+ GM-CSF, TNF $\alpha$  12 days, activated with PMA and ionomycin for 1 hr (D102); DC 70% CD1a+, from CD34+ GM-CSF, TNF $\alpha$  12 days, activated with PMA and ionomycin for 6 hr (D103); DC 95% CD1a+, from CD34+ GM-CSF, TNF $\alpha$  12 days FACS sorted, activated with PMA and  
 35 ionomycin for 1, 6 h pooled (D104); DC 95% CD14+, ex CD34+ GM-CSF, TNF $\alpha$  12 days FACS sorted, activated with PMA and ionomycin 1, 6 hr pooled (D105); DC CD1a+ CD86+, from CD34+ GM-CSF, TNF $\alpha$  12 days FACS sorted, activated with PMA and

ionomycin for 1, 6 h pooled (D106); DC from monocytes GM-CSF, IL-4 5 days, resting (D107); DC from monocytes GM-CSF, IL-4 5 days, resting (D108); DC from monocytes GM-CSF, IL-4 5 days, activated LPS 4, 16 h pooled (D109); DC from monocytes GM-CSF, IL-4 5 days, activated TNF $\alpha$ , monocyte supe for 4, 16 h pooled (D110); leiomyoma L11 benign tumor (X101); normal myometrium M5 (O115); malignant leiomyosarcoma GS1 (X103); lung fibroblast sarcoma line MRC5, activated with PMA and ionomycin for 1, 6 h pooled (C101); kidney epithelial carcinoma cell line CHA, activated with PMA and ionomycin for 1, 6 h pooled (C102); kidney fetal 28 wk male (O100); lung fetal 28 wk male (O101); liver fetal 28 wk male (O102); heart fetal 28 wk male (O103); brain fetal 28 wk male (O104); gallbladder fetal 28 wk male (O106); small intestine fetal 28 wk male (O107); adipose tissue fetal 28 wk male (O108); ovary fetal 25 wk female (O109); uterus fetal 25 wk female (O110); testes fetal 28 wk male (O111); spleen fetal 28 wk male (O112); adult placenta 28 wk (O113); and tonsil inflamed, from 12 year old (X100).

TaqMan quantitative PCR techniques have shown the DCRS6, in both mouse and human, to be expressed on T cells, including thymocytes and CD4+ naive and differentiated (hDCRS6 is also expressed on dendritic cells), in gastrointestinal tissue, including stomach, intestine, colon and associated lymphoid tissue, e.g., Peyer's patches and mesenteric lymph nodes, and upregulated in inflammatory models of bowel disease, e.g., IL-10 KO mice. The hDCRS7 was detected in both resting and activated dendritic cells, epithelial cells, and mucosal tissues, including GI and reproductive tracts. These data suggest that family members are expressed in mucosal tissues and immune system cell types, and/or in gastrointestinal, airway, and reproductive tract development.

As such, therapeutic indications include, e.g., short bowel syndrome, post chemo/radio-therapy or alcoholic recovery, combinations with ulcer treatments or arthritis medication, Th2 pregnancy skewing, stomach lining/tissue regeneration, loss of adsorptive surface conditions, etc. See, e.g., Yamada, et al. (eds. 1999) Textbook of Gastroenterology; Yamada, et al. (eds. 1999) Textbook and Atlas of Gastroenterology; Gore and Levine (2000) Textbook of Gastrointestinal Radiology; and (1987) Textbook of Pediatric Gastroenterology.

Similar samples may isolated in other species for evaluation.

Primers specific for IL-17RA were designed and used in Taqman quantative PCR against various human libraries. IL-17RA is highly expressed in innate immune myeloid cells including dendritic cells and monocytes. Expression is also detected in T-cell libraries. These data demonstrate the receptor is expressed in immune cell types and may be regulated by activation conditions.



Table for IL-17RA library description	CT for IL- 17RA_H
DC ex monocytes GM-CSF, IL-4, resting	16.97
U937 premonocytic line, activated	17.14
DC ex monocytes GM-CSF, IL-4, resting	17.53
DC 70% CD1a+, ex CD34+ GM-CSF, TNFa, resting	18.17
monocytes, LPS, gIFN, anti-IL-10	18.27
DC ex monocytes GM-CSF, IL-4, LPS activated 4+16 hr	18.51
DC ex monocytes GM-CSF, IL-4, monokine activated 4+16 hr	18.68
kidney epithelial carcinoma cell line CHA, activated	18.69
monocytes, LPS, 1 hr	18.72
monocytes, LPS, 6 hr	18.72
DC 70% CD1a+, ex CD34+ GM-CSF, TNFa, activated 1 hr	18.91
DC 70% CD1a+, ex CD34+ GM-CSF, TNFa, activated 6 hr	18.94
T cell, TH1 clone HY06, activated	18.99
lung fetal	19.15
T cell, TH1 clone HY06, resting	19.18
T cell, TH1 clone HY06, anergic	19.23
monocytes, LPS, gIFN, IL-10, 4+16 hr	19.3
spleen fetal	19.51
testes fetal	19.7
T cell, TH0 clone Mot 72, resting	19.71
T cell, TH0 clone Mot 72, resting	19.84
DC CD1a+ CD86+, ex CD34+ GM-CSF, TNFa, activated 1+6 hr	19.94
peripheral blood mononuclear cells, activated	20.01
hematopoietic precursor line TF1, activated	20.07
lung fibroblast sarcoma line MRC5, activated	20.18
Splenocytes, activated	20.21
T cell gd clones, resting	20.27
ovary fetal	20.45
T cells CD4+, TH2 polarized, activated	20.57
Splenocytes, resting	20.6
uterus fetal	20.62
DC 95% CD1a+, ex CD34+ GM-CSF, TNFa, activated 1+6 hr	20.94
epithelial cells, unstimulated	20.96
peripheral blood mononuclear cells, resting	20.97
adipose tissue fetal	21.13

B cell line JY, activated	21.28
monocytes, LPS, gIFN, IL-10	21.37
placenta 28 wk	21.38
NK 20 clones pooled, activated	21.55
pool of two normal human lung samples	21.63
normal human thyroid	21.65
epithelial cells, IL-1b activated	21.72
normal human skin	21.84
T cell, TH0 clone Mot 72, anergic	21.87
small intestine fetal	22.01
CD28- T cell clone in pME	22.08
T cell, TH2 clone HY935, activated	22.09
T cell clones, pooled, resting	22.29
Hashimoto's thyroiditis thyroid sample	22.3
NK 20 clones pooled, resting	22.4
B cell EBV lines, resting	22.45
T cell, TH2 clone HY935, resting	22.86
T cell, TH0 clone Mot 72, activated	23.3
monocytes, LPS, gIFN, anti-IL-10, 4+16 hr	23.39
T cell lines Jurkat and Hut78, resting	23.4
T cell, TH0 clone Mot 72, activated	23.56
<i>Pneumocystis carinii</i> pneumonia lung sample	24.05
U937 premonocytic line, resting	25.01
pool of rheumatoid arthritis samples, human	25.85
pool of three heavy smoker human lung samples	26.1
DC 95% CD14+, ex CD34+ GM-CSF, TNFa, activated 1+6 hr	32.69
kidney fetal	33.7
liver fetal	34.4
NK cytotoxic clone, resting	34.49
tonsil inflamed	35.02
normal w.t. monkey lung	35.45
gallbladder fetal	35.84
TR1 T cell clone	35.86
allergic lung sample	36.39
Psoriasis patient skin sample	36.44
normal human colon	37.34
brain fetal	37.35
<i>Ascaris</i> -challenged monkey lung, 4 hr.	37.75
<i>Ascaris</i> -challenged monkey lung, 24 hr.	40
heart fetal	40
normal w.t. monkey colon	40
ulcerative colitis human colon sample	40

Primers specific for DCRS6\_H were designed and used in Taqman quantitative PCR against various human libraries. DCRS6\_H is expressed in innate immune myeloid cells including dendritic cells and monocytes. Expression is also detected in T-cell libraries. These data demonstrate the receptor is expressed in immune cell types and may be regulated by activation conditions.

5

#### Table for DCRS6\_H

library description	CT for DCRS6_H
T cell, TH0 clone Mot 72, resting	15.54
T cell, TH0 clone Mot 72, resting	15.7
DC ex monocytes GM-CSF, IL-4, resting	17.84
DC ex monocytes GM-CSF, IL-4, resting	18.19
DC ex monocytes GM-CSF, IL-4, LPS activated 4+16 hr	18.3
DC ex monocytes GM-CSF, IL-4, monokine activated 4+16 hr	18.3
T cell, TH1 clone HY06, resting	18.43
NK cytotoxic clone, resting	18.53
T cell clones, pooled, resting	18.8
T cell, TH1 clone HY06, activated	19.03
T cell, TH2 clone HY935, activated	19.1
TR1 T cell clone	19.12
T cells CD4+, TH2 polarized, activated	20.06
B cell EBV lines, resting	20.3
T cell, TH2 clone HY935, resting	20.48
kidney epithelial carcinoma cell line CHA, activated	21.07
T cell, TH1 clone HY06, anergic	21.14
normal human colon	21.29
NK 20 clones pooled, resting	21.49
T cell gd clones, resting	21.58
gallbladder fetal	22.21
kidney fetal	22.79
liver fetal	22.8
<i>Pneumocystis carinii</i> pneumonia lung sample	23.06
CD28- T cell clone in pME	23.18
T cell, TH0 clone Mot 72, anergic	23.2
ovary fetal	23.51
normal human thyroid	24.03
small intestine fetal	24.13
testes fetal	24.82
epithelial cells, IL-1b activated	26.08
pool of three heavy smoker human lung samples	26.49
placenta 28 wk	26.56
normal w.t. monkey lung	28.65
peripheral blood mononuclear cells,	33.39

activated	
<i>Ascaris</i> -challenged monkey lung, 4 hr.	36.59
spleen fetal	38.43
peripheral blood mononuclear cells, resting	40
T cell, TH0 clone Mot 72, activated	40
T cell lines Jurkat and Hut78, resting	40
Splenocytes, resting	40
Splenocytes, activated	40
B cell line JY, activated	40
NK 20 clones pooled, activated	40
hematopoietic precursor line TF1, activated	40
U937 premonocytic line, resting	40
U937 premonocytic line, activated	40
monocytes, LPS, gIFN, anti-IL-10	40
monocytes, LPS, gIFN, IL-10	40
monocytes, LPS, gIFN, anti-IL-10, 4+16 hr	40
monocytes, LPS, gIFN, IL-10, 4+16 hr	40
monocytes, LPS, 1 hr	40
monocytes, LPS, 6 hr	40
DC 70% CD1a+, ex CD34+ GM-CSF, TNFa, resting	40
DC 70% CD1a+, ex CD34+ GM-CSF, TNFa, activated 1 hr	40
DC 70% CD1a+, ex CD34+ GM-CSF, TNFa, activated 6 hr	40
DC 95% CD1a+, ex CD34+ GM-CSF, TNFa, activated 1+6 hr	40
DC 95% CD14+, ex CD34+ GM-CSF, TNFa, activated 1+6 hr	40
DC CD1a+ CD86+, ex CD34+ GM-CSF, TNFa, activated 1+6 hr	40
epithelial cells, unstimulated	40
lung fibroblast sarcoma line MRC5, activated	40
<i>Ascaris</i> -challenged monkey lung, 24 hr.	40
pool of two normal human lung samples	40
allergic lung sample	40
normal w.t. monkey colon	40
ulcerative colitis human colon sample	40
Hashimoto's thyroiditis thyroid sample	40
pool of rheumatoid arthritis samples, human	40
normal human skin	40
Psoriasis patient skin sample	40
tonsil inflamed	40
lung fetal	40
heart fetal	40
brain fetal	40
adipose tissue fetal	40
uterus fetal	40

T cell, TH0 clone Mot 72, activated 40

- 5 Primers specific for DCRS7\_H were designed and used in Taqman quantitative PCR against various human libraries. DCRS7\_H is expressed in innate immune myeloid cells including dendritic cells and monocytes. Expression is also detected in fetal libraries. These data demonstrate the receptor is expressed in immune cell types and may be regulated by activation conditions.

library description	CT for DCRS7_H
fetal uterus	19.05
DC mix	19.34
fetal small intestine	19.46
fetal ovary	19.68
fetal testes	19.75
fetal lung	20.04
CHA	20.24
normal thyroid	20.32
DC/GM/IL-4	20.52
fetal spleen	20.86
normal lung	20.94
TF1	21
allergic lung #19	21.02
Psoriasis skin	21.07
fetal liver	21.15
MRC5	21.15
24 hr. Ascaris lung	21.17
hi dose IL-4 lung	21.23
CD1a+ 95%	21.32
Hashimotos thyroiditis	21.35
Crohns colon 4003197A	21.35
normal lung pool	21.36
70% DC resting	21.42
fetal kidney	21.58
adult placenta	21.68
lung 121897-1	21.8
Pneumocystis carinii lung	21.81
#20	
A549 unstim.	21.89
normal colon #22	21.94
18 hr. Ascaris lung	22.09
normal skin	22.1
Crohns colon 9609C144	22.13
fetal adipose tissue	22.35
D6	22.39

DC resting CD34-derived	22.45
DC TNF/TGFb act CD34-der.	22.54
fetal brain	22.9
DC CD40L activ. mono-deriv.	22.91
Crohns colon 403242A	22.91
ulcerative colitis colon #26	23
RA synovium pool	23.06
A549 activated	23.06
mono + IL-10	23.42
DC LPS	23.49
Mot 72 activated	23.66
CD1a+ CD86+	23.86
HY06 resting	23.87
U937 activated	23.97
inflammed tonsil	23.97
D1	24.06
M1	24.17
CD14+ 95%	24.21
lung 080698-2	24.28
4 hr. Ascaris lung	24.37
Jurkat activated pSPORT	24.42
DC resting mono-derived	24.48
HY06 activated	24.54
C+	24.64
Splenocytes resting	24.65
U937/CD004 resting	24.96
PBMC resting	25.8
Mot 72 resting	25.91
mono + anti-IL-10	26.14
NK pool	26.99
HY06 anti-peptide	27.34
mast cell pME	27.38
Tc gamma delta	28.14
TC1080 CD28- pMET7	31.05
PBMC activated	31.89
NK non cytotox.	32.3
RV-C30 TR1 pMET7	32.5
Bc	33.72
C-	33.8
Splenocytes activated	34.7
JY	35.05
NK cytotox.	36.44
NKL/IL-2	37.59
HY935 resting	37.6
NK pool activated	38.15
Mot 72 anti-peptide	38.87
fetal heart	40.92

B21 resting	42.05
Jurkat resting pSPORT	42.8
B21 activated	43.09
NKA6 pSPORT	44.85
HY935 activated	45
M6	45

- 5 Primers specific for DCRS9\_H were designed and used in Taqman quantitative PCR against various human libraries. DCRS9\_H is expressed T-cells, fetal lung, and resting monocytes. These data demonstrate the receptor is expressed in immune cell types and may be regulated by activation conditions.

Table for DCRS9\_H  
library description CT for  
DCRS9\_H

HY06 resting	22.35
fetal lung	22.63
HY06 anti-peptide	22.72
HY06 activated	22.96
U937/CD004 resting	24.16
fetal small intestine	24.94
JY	25.04
Mot 72 resting	25.12
Jurkat activated pSPORT	25.2
RV-C30 TR1 pMET7	26.51
fetal kidney	26.76
MRC5	27.2
Psoriasis skin	27.3
Tc gamma delta	27.37
Crohns colon 4003197A	27.44
fetal spleen	27.72
normal lung	27.83
Hashimotos thyroiditis	28.03
B21 resting	28.32
TF1	28.39
NK cytotox.	28.44
TC1080 CD28- pMET7	28.61
Pneumocystis carinii lung #20	29.05
U937 activated	29.06
HY935 resting	29.09
CD1a+ 95%	29.13

B21 activated	29.2
Mot 72 activated	29.21
fetal testes	29.27
lung 080698-2	29.32
Jurkat resting	29.38
pSPORT	
CD14+ 95%	29.38
normal thyroid	29.53
Mot 72 anti-peptide	29.65
Splenocytes resting	29.85
Crohns colon 9609C144	30.28
lung 121897-1	30.37
24 hr. Ascaris lung	30.59
hi dose IL-4 lung	30.8
CD1a+ CD86+	31.42
normal skin	31.73
fetal uterus	31.79
PBMC activated	31.82
inflammed tonsil	31.98
fetal brain	32.21
RA synovium pool	32.77
allergic lung #19	33.18
18 hr. Ascaris lung	33.42
adult placenta	33.43
normal lung pool	33.45
Crohns colon 403242A	33.52
NK pool	33.72
HY935 activated	33.75
DC/GM/IL-4	34.28
DC resting mono-derived	34.57
fetal ovary	35.06
fetal adipose tissue	35.07
CHA	35.2
PBMC resting	35.95
Bc	36.19
A549 unstim.	36.4
fetal heart	36.87
ulcerative colitis colon #26	37.83
C-	38.32
4 hr. Ascaris lung	40.2
D6	40.62
C+	44.38



A549 activated	44.58
Splenocytes	45
activated	
NK pool activated	45
NKA6 pSPORT	45
NKL/IL-2	45
NK non cytotox.	45
mono + anti-IL-10	45
mono + IL-10	45
M1	45
M6	45
70% DC resting	45
D1	45
DC LPS	45
DC mix	45
fetal liver	45
mast cell pME	45
DC CD40L activ.	45
mono-deriv.	
DC resting CD34- derived	45
DC TNF/TGFb act	45
CD34-der.	
normal colon #22	45

## V. Cloning of species counterparts

Various strategies are used to obtain species counterparts of the DCRSs, preferably from other primates or rodents. One method is by cross hybridization using closely related species DNA probes. It may be useful to go into evolutionarily similar species as intermediate steps. Another method is by using specific PCR primers based on the identification of blocks of similarity or difference between genes, e.g., areas of highly conserved or nonconserved polypeptide or nucleotide sequence. Sequence database searches may identify species counterparts.

## 10 VI. Production of mammalian protein

An appropriate, e.g., GST, fusion construct is engineered for expression, e.g., in *E. coli*. For example, a mouse IGIF pGex plasmid is constructed and transformed into *E. coli*. Freshly transformed cells are grown, e.g., in LB medium containing 50 µg/ml ampicillin and induced with IPTG (Sigma, St. Louis, MO). After overnight induction, the bacteria are harvested and the pellets containing the appropriate protein are isolated. The pellets are homogenized, e.g., in TE buffer (50 mM Tris-base pH 8.0, 10 mM EDTA and 2 mM pefabloc) in 2 liters. This material is passed through a microfluidizer (Microfluidics, Newton, MA) three times. The fluidized supernatant is spun down on a Sorvall GS-3 rotor for 1 h at 13,000 rpm. The resulting supernatant containing the cytokine receptor protein is filtered and passed over a glutathione-SEPHAROSE column equilibrated in 50 mM Tris-base pH 8.0. Fractions containing the DCRS8-GST fusion protein are pooled and cleaved, e.g., with thrombin (Enzyme Research Laboratories, Inc., South Bend, IN). The cleaved pool is then passed over a Q-SEPHAROSE column equilibrated in 50 mM Tris-base. Fractions containing DCRS8 are pooled and diluted in cold distilled H<sub>2</sub>O, to lower the conductivity, and passed back over a fresh Q-Sepharose column, alone or in succession with an immunoaffinity antibody column. Fractions containing the DCRS8 protein are pooled, aliquoted, and stored in the -70° C freezer.

Comparison of the CD spectrum with cytokine receptor protein may suggest that the protein is correctly folded. See Hazuda, et al. (1969) J. Biol. Chem. 264:1689-1693.

## 30 VII. Preparation of specific antibodies

Inbred Balb/c mice are immunized intraperitoneally with recombinant forms of the protein, e.g., purified DCRS8 or stable transfected NIH-3T3 cells. Animals are boosted at appropriate time points with protein, with or without additional adjuvant, to further stimulate antibody production. Serum is collected, or hybridomas produced with harvested spleens.

Alternatively, Balb/c mice are immunized with cells transformed with the gene or fragments thereof, either endogenous or exogenous cells, or with isolated membranes enriched for expression of the antigen. Serum is collected at the appropriate time, typically after numerous further administrations. Various gene therapy techniques may be useful, e.g., in producing protein in situ, for generating an immune response. Serum or antibody preparations may be cross-absorbed or immunoselected to prepare substantially purified antibodies of defined specificity and high affinity.

Monoclonal antibodies may be made. For example, splenocytes are fused with an appropriate fusion partner and hybridomas are selected in growth medium by standard procedures. Hybridoma supernatants are screened for the presence of antibodies which bind to the DCRS8, e.g., by ELISA or other assay. Antibodies which specifically recognize specific DCRS8 embodiments may also be selected or prepared.

In another method, synthetic peptides or purified protein are presented to an immune system to generate monoclonal or polyclonal antibodies. See, e.g., Coligan (ed. 1991) Current Protocols in Immunology Wiley/Greene; and Harlow and Lane (1989) Antibodies: A Laboratory Manual Cold Spring Harbor Press. In appropriate situations, the binding reagent is either labeled as described above, e.g., fluorescence or otherwise, or immobilized to a substrate for panning methods. Nucleic acids may also be introduced into cells in an animal to produce the antigen, which serves to elicit an immune response. See, e.g., Wang, et al. (1993) Proc. Nat'l. Acad. Sci. 90:4156-4160; Barry, et al. (1994) BioTechniques 16:616-619; and Xiang, et al. (1995) Immunity 2: 129-135.

#### VIII. Production of fusion proteins

Various fusion constructs are made with DCRS8 or DCRS9. A portion of the appropriate gene is fused to an epitope tag, e.g., a FLAG tag, or to a two hybrid system construct. See, e.g., Fields and Song (1989) Nature 340:245-246.

The epitope tag may be used in an expression cloning procedure with detection with anti-FLAG antibodies to detect a binding partner, e.g., ligand for the respective cytokine receptor. The two hybrid system may also be used to isolate proteins which specifically bind to the receptor subunit.

#### IX. Structure activity relationship

Information on the criticality of particular residues is determined using standard procedures and analysis. Standard mutagenesis analysis is performed, e.g., by generating many different variants at determined positions, e.g., at the positions identified above, and evaluating biological activities of the variants. This may be performed to the extent of determining positions which modify activity, or to focus on specific positions to

determine the residues which can be substituted to either retain, block, or modulate biological activity.

Alternatively, analysis of natural variants can indicate what positions tolerate natural mutations. This may result from populational analysis of variation among  
5 individuals, or across strains or species. Samples from selected individuals are analyzed, e.g., by PCR analysis and sequencing. This allows evaluation of population polymorphisms.

#### X. Isolation of a ligand

10 A cytokine receptor can be used as a specific binding reagent to identify its binding partner, by taking advantage of its specificity of binding, much like an antibody would be used. The binding receptor may be a heterodimer of receptor subunits; or may involve, e.g., a complex of the DCRS8 with another cytokine receptor subunit. A binding reagent is either labeled as described above, e.g., fluorescence or otherwise, or  
15 immobilized to a substrate for panning methods.

The binding composition is used to screen an expression library made from a cell line which expresses a binding partner, i.e., ligand, preferably membrane associated. Standard staining techniques are used to detect or sort surface expressed ligand, or surface expressing transformed cells are screened by panning. Screening of intracellular  
20 expression is performed by various staining or immunofluorescence procedures. See also McMahan, et al. (1991) EMBO J. 10:2821-2832.

For example, on day 0, precoat 2-chamber permanox slides with 1 ml per chamber of fibronectin, 10 ng/ml in PBS, for 30 min at room temperature. Rinse once with PBS. Then plate COS cells at  $2-3 \times 10^5$  cells per chamber in 1.5 ml of growth media. Incubate  
25 overnight at 37 C.

On day 1 for each sample, prepare 0.5 ml of a solution of 66  $\mu$ g/ml DEAE-dextran, 66  $\mu$ M chloroquine, and 4  $\mu$ g DNA in serum free DME. For each set, a positive control is prepared, e.g., of DCRS8-FLAG cDNA at 1 and 1/200 dilution, and a negative mock. Rinse cells with serum free DME. Add the DNA solution and incubate 5 hr at 37  
30 C. Remove the medium and add 0.5 ml 10% DMSO in DME for 2.5 min. Remove and wash once with DME. Add 1.5 ml growth medium and incubate overnight.

On day 2, change the medium. On days 3 or 4, the cells are fixed and stained. Rinse the cells twice with Hank's Buffered Saline Solution (HBSS) and fix in 4% paraformaldehyde (PFA)/glucose for 5 min. Wash 3X with HBSS. The slides may be  
35 stored at -80 C after all liquid is removed. For each chamber, 0.5 ml incubations are performed as follows. Add HBSS/saponin (0.1%) with 32  $\mu$ l/ml of 1 M NaN<sub>3</sub> for 20 min. Cells are then washed with HBSS/saponin 1X. Add appropriate DCRS8 or

DCRS8/antibody complex to cells and incubate for 30 min. Wash cells twice with HBSS/saponin. If appropriate, add first antibody for 30 min. Add second antibody, e.g., Vector anti-mouse antibody, at 1/200 dilution, and incubate for 30 min. Prepare ELISA solution, e.g., Vector Elite ABC horseradish peroxidase solution, and preincubate for 30 min. Use, e.g., 1 drop of solution A (avidin) and 1 drop solution B (biotin) per 2.5 ml HBSS/saponin. Wash cells twice with HBSS/saponin. Add ABC HRP solution and incubate for 30 min. Wash cells twice with HBSS, second wash for 2 min, which closes cells. Then add Vector diaminobenzoic acid (DAB) for 5 to 10 min. Use 2 drops of buffer plus 4 drops DAB plus 2 drops of H<sub>2</sub>O<sub>2</sub> per 5 ml of glass distilled water.

Carefully remove chamber and rinse slide in water. Air dry for a few minutes, then add 1 drop of Crystal Mount and a cover slip. Bake for 5 min at 85-90 C.

Evaluate positive staining of pools and progressively subclone to isolation of single genes responsible for the binding.

Alternatively, receptor reagents are used to affinity purify or sort out cells expressing a putative ligand. See, e.g., Sambrook, et al. or Ausubel, et al.

Another strategy is to screen for a membrane bound receptor by panning. The receptor cDNA is constructed as described above. Immobilization may be achieved by use of appropriate antibodies which recognize, e.g., a FLAG sequence of a DCRS8 fusion construct, or by use of antibodies raised against the first antibodies. Recursive cycles of selection and amplification lead to enrichment of appropriate clones and eventual isolation of receptor expressing clones.

Phage expression libraries can be screened by mammalian DCRS8. Appropriate label techniques, e.g., anti-FLAG antibodies, will allow specific labeling of appropriate clones.

We tested the ability of DCRS receptors to specifically bind IL-17 family cytokines. Recombinant FLAG-hIL-17 family cytokines were used in binding experiments on Baf/3 DCRS receptor transfected expressing recombinant IL-17R\_H, DCRS6\_H, DCRS7\_H, DCRS8\_H and DCRS9\_H and analyzed by FACS. We can demonstrate specific binding of IL-17 family member IL-74 to DCRS6 expressing Baf/3 cells. In additional experiments we have shown IL-17 specific binding to IL-17R\_H, DCRS7\_H, DCRS8\_H. Further experiments show IL-71 binding to DCRS8\_Hu transfectants. These experiments demonstrate the sequence homology among IL-17 related cytokine receptors confers functional binding to IL-17 cytokines.

All citations herein are incorporated herein by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Many modifications and variations of this invention can be made without departing from its spirit and scope, as will be apparent to those skilled in the art. The specific embodiments described herein are offered by way of example only, and the invention is to be limited by the terms of the appended claims, along with the full scope  
5 of equivalents to which such claims are entitled; and the invention is not to be limited by the specific embodiments that have been presented herein by way of example.

## WHAT IS CLAIMED IS:

1. A composition of matter selected from:
  - a) a substantially pure or recombinant polypeptide comprising at least three distinct nonoverlapping segments of at least four amino acids identical to segments of SEQ ID NO: 14;
  - b) a substantially pure or recombinant polypeptide comprising at least two distinct nonoverlapping segments of at least five amino acids identical to segments of SEQ ID NO: 14;
  - c) a natural sequence DCRS8 comprising mature SEQ ID NO: 14;
  - d) a fusion polypeptide comprising DCRS8 sequence;
  - e) a substantially pure or recombinant polypeptide comprising at least three distinct nonoverlapping segments of at least four amino acids identical to segments of SEQ ID NO: 17 or 20;
  - f) a substantially pure or recombinant polypeptide comprising at least two distinct nonoverlapping segments of at least five amino acids identical to segments of SEQ ID NO: 17 or 20;
  - g) a natural sequence DCRS9 comprising mature SEQ ID NO: 17 or 20; or
  - h) a fusion polypeptide comprising DCRS9 sequence.
2. The substantially pure or isolated antigenic polypeptide of Claim 1, wherein said distinct nonoverlapping segments of identity include:
  - a) one of at least eight amino acids;
  - b) one of at least four amino acids and a second of at least five amino acids;
  - c) at least three segments of at least four, five, and six amino acids, or
  - d) one of at least twelve amino acids.
3. The composition of matter of Claim 1, wherein said:
  - a) polypeptide:
    - i) comprises a mature sequence of Table 3 or 4;
    - ii) is an unglycosylated form of DCRS8 or DCRS9;
    - iii) is from a primate, such as a human;
    - iv) comprises at least seventeen amino acids of SEQ ID NO: 14 or 17;
    - v) exhibits at least four nonoverlapping segments of at least seven amino acids of SEQ ID NO: 14 or 17;
    - vi) is a natural allelic variant of DCRS8 or DCRS9;
    - vii) has a length at least about 30 amino acids;

- viii) exhibits at least two non-overlapping epitopes which are specific for a primate DCRS8 or DCRS9;
- ix) is glycosylated;
- x) has a molecular weight of at least 30 kD with natural glycosylation;
- xi) is a synthetic polypeptide;
- xii) is attached to a solid substrate;
- xiii) is conjugated to another chemical moiety;
- xiv) is a 5-fold or less substitution from natural sequence; or
- xv) is a deletion or insertion variant from a natural sequence.
4. A composition comprising:
- a) a substantially pure DCRS8 or DCRS9 and another cytokine receptor family member;
- b) a sterile DCRS8 or DCRS9 polypeptide of Claim 1;
- c) said DCRS8 or DCRS9 polypeptide of Claim 1 and a carrier, wherein said carrier is:
- i) an aqueous compound, including water, saline, and/or buffer; and/or
- ii) formulated for oral, rectal, nasal, topical, or parenteral administration.
5. The fusion polypeptide of Claim 1, comprising:
- a) mature protein sequence of Table 3 or 4;
- b) a detection or purification tag, including a FLAG, His6, or Ig sequence; or
- c) sequence of another cytokine receptor protein.
6. A kit comprising a polypeptide of Claim 1, and:
- a) a compartment comprising said protein or polypeptide; or
- b) instructions for use or disposal of reagents in said kit.
7. A binding compound comprising an antigen binding site from an antibody, which specifically binds to a natural DCRS8 or DCRS9 polypeptide of Claim 1, wherein:
- a) said binding compound is in a container;
- b) said DCRS8 or DCRS9 polypeptide is from a human;
- c) said binding compound is an Fv, Fab, or Fab2 fragment;
- d) said binding compound is conjugated to another chemical moiety; or
- e) said antibody:
- i) is raised against a peptide sequence of a mature polypeptide of Table 3 or 4;



- ii) is raised against a mature DCRS8 or DCRS9;  
iii) is raised to a purified human DCRS8 or DCRS9;  
iv) is immunoselected;  
v) is a polyclonal antibody;  
5 vi) binds to a denatured DCRS8 or DCRS9;  
vii) exhibits a Kd to antigen of at least 30  $\mu$ M;  
viii) is attached to a solid substrate, including a bead or plastic membrane;  
ix) is in a sterile composition; or  
x) is detectably labeled, including a radioactive or fluorescent label.
- 10 8. A kit comprising said binding compound of Claim 7, and:  
a) a compartment comprising said binding compound; or  
b) instructions for use or disposal of reagents in said kit.
- 15 9. A method of producing an antigen:antibody complex, comprising  
contacting under appropriate conditions a primate DCRS8 or DCRS9 polypeptide with an  
antibody of Claim 7, thereby allowing said complex to form.
- 20 10. The method of Claim 9, wherein:  
a) said complex is purified from other cytokine receptors;  
b) said complex is purified from other antibody;  
c) said contacting is with a sample comprising an interferon;  
d) said contacting allows quantitative detection of said antigen;  
e) said contacting is with a sample comprising said antibody; or  
25 f) said contacting allows quantitative detection of said antibody.
11. A composition comprising:  
a) a sterile binding compound of Claim 7, or  
b) said binding compound of Claim 7 and a carrier, wherein said carrier is:  
30 i) an aqueous compound, including water, saline, and/or buffer; and/or  
ii) formulated for oral, rectal, nasal, topical, or parenteral administration.
12. An isolated or recombinant nucleic acid encoding said polypeptide of  
Claim 1, wherein said:  
35 a) DCRS8 or DCRS9 is from a human; or  
b) said nucleic acid:  
i) encodes an antigenic peptide sequence of Table 3 or 4;

- ii) encodes a plurality of antigenic peptide sequences of Table 3 or 4;  
iii) exhibits identity over at least thirteen nucleotides to a natural cDNA  
encoding said segment;  
iv) is an expression vector;  
5 v) further comprises an origin of replication;  
vi) is from a natural source;  
vii) comprises a detectable label;  
viii) comprises synthetic nucleotide sequence;  
ix) is less than 6 kb, preferably less than 3 kb;  
10 x) is from a primate;  
xi) comprises a natural full length coding sequence;  
xii) is a hybridization probe for a gene encoding said DCRS8 or DCRS9;  
or  
xiii) is a PCR primer, PCR product, or mutagenesis primer.
- 15 13. A cell or tissue comprising said recombinant nucleic acid of Claim 12.
14. The cell of Claim 13, wherein said cell is:  
a) a prokaryotic cell;  
20 b) a eukaryotic cell;  
c) a bacterial cell;  
d) a yeast cell;  
e) an insect cell;  
f) a mammalian cell;  
25 g) a mouse cell;  
h) a primate cell; or  
i) a human cell.
15. A kit comprising said nucleic acid of Claim 12, and:  
30 a) a compartment comprising said nucleic acid;  
b) a compartment further comprising a primate DCRS8 or DCRS9 polypeptide;  
or  
c) instructions for use or disposal of reagents in said kit.
- 35 16. A nucleic acid which:  
a) hybridizes under wash conditions of 30 minutes at 30° C and less than 2M salt  
to the coding portion of SEQ ID NO: 13 or 16; or

- b) exhibits identity over a stretch of at least about 30 nucleotides to a primate DCRS8 or DCRS9.

17. The nucleic acid of Claim 16, wherein:

- 5 a) said wash conditions are at 45° C and/or 500 mM salt; or  
b) said stretch is at least 55 nucleotides.

18. The nucleic acid of Claim 16, wherein:

- 10 a) said wash conditions are at 55° C and/or 150 mM salt; or  
b) said stretch is at least 75 nucleotides.

19. A method of modulating physiology or development of a cell or tissue culture cells comprising contacting said cell with an agonist or antagonist of a mammalian DCRS8 or DCRS9.

15

20. The method of Claim 19, wherein said cell is transformed with a nucleic acid encoding said DCRS8 or DCRS9 and another cytokine receptor subunit.

DCRS7_Mu	RTALLHSADG-AGYERLVGALASALSQMP-- --LRVAVDLWSRRE-LSAHGALAWFHHQR
DCRS7_Hu	RAALLYSADD-SGFERLVGALASALCQLP-- --LRVAVDLWSRRE-LSAQGPVAFWFAQR
IL-17R_Hu	RKVWIIYSADH-PLYVDVLKFAQFLLTACG-- --TEVALDLLEEQA-ISEAGVMTWVGRQK
IL-17R_Mu	RKVWIVYSADH-PLYVEVLKFAQFLITACG-- --TEVALDLLEEQV-ISEVGVMTWVSRQK
DCRS10	RKVFITYSMD-- --TAMEVVVKFVNFLLVNG-- --FQTAIDIFEDR-- --IRGIDI IKWMERYL
DCRS10_Mu	RKVFITYSMD-- --TAMEVVVKFVNFLLVNG-- --FQTAIDIFEDR-- --IRGIDI IKWMERYL
DCRS9_Hu	RPVLLHAADS-EAQRRLVGALAE LLRAALGGRDVIVDLWEGRH-VARVGPLPWLWAAR
DCRS8_Hu	PKVFLCYSSKDGQNHMNVVQCFA YFLQDFCG-- --CEVALDLWEDFS- LCREGQREWVIQKI
IL-17R_Ce	VKVMIVYADDN-DLHTDCVKKLVENLRNCAS-- --CDPVFDLEKLI-- --TAEIVPSRWLVDQI
DCRS6_Hu	IKVLVVPSEI-- --CFHHTICYFTEFLQNHCR-- --SEVILEKWQKKK-IAEMGPVQWLATQK
DCRS6_Ce	FKVMLVCPEVS-GRDEDFMMRIADALKKSN-- --NKVVCDRFWEDSKNAEENMLHWVYEQT

DCRS7_Mu	RRILQEGGVILLFSPAAVAQCQ----
DCRS7_Hu	RQTLQEGGVVLLFSPGAVALCS----EWLQDGVSGPGAHP----HDAFRASLSCVLPDPFL
IL-17R_Hu	QEMVESNSKIIVLCSTRGTRAKWQALLGRGAP-VRLRCDHGKPV-GDLFTAAMNMILPDPFK
IL-17R_Mu	QEMVESNSKIIILCSRGTQAKWKAILGWAEPVQLRCDHWKPA-GDLFTAAMNMILPDPFK
DCRS10	R---DKTVMIIVAISPKYKQDVE-----GAESQLED- EHGL---HTKYIHRM-MQIEFIK
DCRS10_Mu	R---DKTVMIIVAISPKYKQDVE-----GAESQLED- EHGL---HTKYIHRM-MQIEFIS
DCRS9_Hu	TRVAREQGTVLLWSGADLRPVS-----GDPD-RAAP-----LLA-----LLHAAP
DCRS8_Hu	H----ESQFIIVVCSKGMKYFVD----KKNYKHKGGRGSGK----GELFLVAVSAIAEKLRL
IL-17R_Ce	S----SLKKFIIIVSDCAEKILD-----TEASETHQLVQARP---FADLFGPAMENIIRDAT
DCRS6_Hu	K----AADKVVFLLSNDVNSVCD-----GTCGKSEGSPSENS---QDLFPLAFNLFCSDLRL
DCRS6_Ce	K----IAEKIIVFHSAYYHPRCG---IYDVINNFFPCTDPR-----LAHIALT---PEAQ

**FIG. 1A**

DCRS7_Mu	QGRATGR-----YVGVFYDGLLHPDSVPSFVRVAPLFSLP-SQLPAFLDALQ--GGCSTS
DCRS7_Hu	QGRAPGS-----YVGACFDRLLHPDAVPALFRTVPVFTLP-SQLPDFLGALQ--QPRAPR
IL-17R_Hu	RPACFGT-----YVVCYFSEVSCDGDVPDLFGAAPRYPLM-DRFEEVYFRIQ--DLEMFO
IL-17R_Mu	RPACFGT-----YVVCYFSGICSERDVPDLFNITSRYPLM-DRFEEVYFRIQ--DLEMFE
DCRS10	QGS MNFR-----FIPVLF PNAK-KEHVPTWLQNT HVYSWP-KNKKNILLRLL-REEEYVA
DCRS10_Mu	QGS MNFR-----FIPVLF PNAK-KEHVPTWLQNT HVYSWP-KNKKNILLRLL-REEEYVA
DCRS9_Hu	RPL-----LLLAYFSRLCAKGDIPPLRALPRYRLL-RDLPRLLRALD--ARPPFAE
DCRS8_Hu	QAKQSSAALSKFIAVYFDYSC-EGDVPGILDLSTKYRLM-DNLPQLCSHLHSRDHGLQE
IL-17R_Ce	HNFPEAR---KKYAVVRFNYS P---HVPPNLA I LNLPTFIPEQFAQLTAF LHN-VEHTER
DCRS6_Hu	SQIHLHK-----YVVVYFREID-TKDDYNALSVCPKYHLM-KDATAFCAELL---HVKQQ
DCRS6_Ce	RSVPKEV---EYVLPRDQKLL--EDAFDITIADPLVIDIPIEDVAIPENVP--IHHESC

;

DCRS7_Mu	AGRPAD RVER-----VT-----QALRSALDSCTS-----
DCRS7_Hu	SGRLQERAEQ-----VS-----RALQPALDSYFHPP-----
IL-17R_Hu	PGRMHRV GELSGDNYLRS---PGRQLRAALDRFRDWQVRC PDW
IL-17R_Mu	PGRMHHVREL TGDNYLQS---PSGRQLKEAVLRFQEWQTCCPDW
DCRS10	P-----PRG PL-----PTLQVVPL-----
DCRS10_Mu	P-----PRG PL-----PTLQVVPL-----
DCRS9_Hu	ATSWGRLGAR-----QRRQSRLELCSR-----
DCRS8_Hu	PGQHTRQGS R-----RNYFRSKSGRSLYVAICNMHQFIDE EPDW
IL-17R_Ce	ANVTQNISEA-----Q-----IHEWNLCASRMMSFFVRNP NW
DCRS6_Hu	VS-----AGKR-----SQACHDGCCSL-----
DCRS6_Ce	DSIDSRNNSK-----THSTD SGVSSLS S-----NS--

:

FIG. 1B

## SEQUENCE SUBMISSION

SEQ ID NO: 1 is primate DCRS6 nucleotide sequence.  
SEQ ID NO: 2 is primate DCRS6 polypeptide sequence.  
SEQ ID NO: 3 is primate DCRS6 reverse translation.  
SEQ ID NO: 4 is rodent DCRS6 nucleotide sequence.  
SEQ ID NO: 5 is rodent DCRS6 polypeptide sequence.  
SEQ ID NO: 6 is rodent DCRS6 reverse translation.  
SEQ ID NO: 7 is primate DCRS7 nucleotide sequence.  
SEQ ID NO: 8 is primate DCRS7 polypeptide sequence.  
SEQ ID NO: 9 is primate DCRS7 reverse translation.  
SEQ ID NO: 10 is rodent DCRS7 nucleotide sequence.  
SEQ ID NO: 11 is rodent DCRS7 polypeptide sequence.  
SEQ ID NO: 12 is rodent DCRS7 reverse translation.  
SEQ ID NO: 13 is primate DCRS8 nucleotide sequence.  
SEQ ID NO: 14 is primate DCRS8 polypeptide sequence.  
SEQ ID NO: 15 is primate DCRS8 reverse translation.  
SEQ ID NO: 16 is primate DCRS9 nucleotide sequence.  
SEQ ID NO: 17 is primate DCRS9 polypeptide sequence.  
SEQ ID NO: 18 is primate DCRS9 reverse translation.  
SEQ ID NO: 19 is rodent DCRS9 nucleotide sequence.  
SEQ ID NO: 20 is rodent DCRS9 polypeptide sequence.  
SEQ ID NO: 21 is rodent DCRS9 reverse translation.  
SEQ ID NO: 22 is primate DCRS10 nucleotide sequence.  
SEQ ID NO: 23 is primate DCRS10 polypeptide sequence.  
SEQ ID NO: 24 is primate DCRS10 reverse translation.  
SEQ ID NO: 25 is rodent DCRS10 nucleotide sequence.  
SEQ ID NO: 26 is rodent DCRS10 polypeptide sequence.  
SEQ ID NO: 27 is rodent DCRS10 reverse translation.  
SEQ ID NO: 28 is primate IL-17 receptor peptide sequence.  
SEQ ID NO: 29 is rodent IL-17 receptor peptide sequence.  
SEQ ID NO: 30 is worm IL-17 receptor peptide sequence.  
SEQ ID NO: 31 is worm DCRS6 nucleotide sequence.

<110> Schering Corporation

<120> Mammalian Receptor Proteins; Related Reagents and  
Methods

<130> DX01170K PCT

<140>

<141>

<150> US 60/206,862

<151> 2000-05-24

<160> 31

<170> PatentIn Ver. 2.0

<210> 1

<211> 1796

<212> DNA

<213> Unknown

<220>

<223> Description of Unknown Organism: primate; surmised  
Homo sapiens

<220>  
 <221> CDS  
 <222> (4)..(1509)

<220>  
 <221> mat\_peptide  
 <222> (46)..(1509)

```

<400> 1
gcg atg tcg ctc gtg ctg cta agc ctg gcc gcg ctg tgc agg agc gcc 48
    Met Ser Leu Val Leu Leu Ser Leu Ala Ala Leu Cys Arg Ser Ala
                    -10                    -5                    -1    1

gta ccc cga gag ccg acc gtt caa tgt ggc tct gaa act ggg cca tct 96
Val Pro Arg Glu Pro Thr Val Gln Cys Gly Ser Glu Thr Gly Pro Ser
                    5                    10                    15

cca gag tgg atg cta caa cat gat cta atc ccg gga gac ttg agg gac 144
Pro Glu Trp Met Leu Gln His Asp Leu Ile Pro Gly Asp Leu Arg Asp
                    20                    25                    30

ctc cga gta gaa cct gtt aca act agt gtt gca aca ggg gac tat tca 192
Leu Arg Val Glu Pro Val Thr Thr Ser Val Ala Thr Gly Asp Tyr Ser
                    35                    40                    45

att ttg atg aat gta agc tgg gta ctc cgg gca gat gcc agc atc cgc 240
Ile Leu Met Asn Val Ser Trp Val Leu Arg Ala Asp Ala Ser Ile Arg
                    50                    55                    60                    65

ttg ttg aag gcc acc aag att tgt gtg acg ggc aaa agc aac ttc cag 288
Leu Leu Lys Ala Thr Lys Ile Cys Val Thr Gly Lys Ser Asn Phe Gln
                    70                    75                    80

tcc tac agc tgt gtg agg tgc aat tac aca gag gcc ttc cag act cag 336
Ser Tyr Ser Cys Val Arg Cys Asn Tyr Thr Glu Ala Phe Gln Thr Gln
                    85                    90                    95

acc aga ccc tct ggt ggt aaa tgg aca ttt tcc tat atc ggc ttc cct 384
Thr Arg Pro Ser Gly Gly Lys Trp Thr Phe Ser Tyr Ile Gly Phe Pro
                    100                    105                    110

gta gag ctg aac aca gtc tat ttc att ggg gcc cat aat att cct aat 432
Val Glu Leu Asn Thr Val Tyr Phe Ile Gly Ala His Asn Ile Pro Asn
                    115                    120                    125

gca aat atg aat gaa gat ggc cct tcc atg tct gtg aat ttc acc tca 480
Ala Asn Met Asn Glu Asp Gly Pro Ser Met Ser Val Asn Phe Thr Ser
                    130                    135                    140                    145

cca ggc tgc cta gac cac ata atg aaa tat aaa aaa aag tgt gtc aag 528
Pro Gly Cys Leu Asp His Ile Met Lys Tyr Lys Lys Lys Cys Val Lys
                    150                    155                    160

gcc gga agc ctg tgg gat ccg aac atc act gct tgt aag aag aat gag 576
Ala Gly Ser Leu Trp Asp Pro Asn Ile Thr Ala Cys Lys Lys Asn Glu
                    165                    170                    175

gag aca gta gaa gtg aac ttc aca acc act ccc ctg gga aac aga tac 624
Glu Thr Val Glu Val Asn Phe Thr Thr Thr Pro Leu Gly Asn Arg Tyr
                    180                    185                    190

```

atg gct ctt atc caa cac agc act atc atc ggg ttt tct cag gtg ttt	672
Met Ala Leu Ile Gln His Ser Thr Ile Ile Gly Phe Ser Gln Val Phe	
195 200 205	
gag cca cac cag aag aaa caa acg cga gct tca gtg gtg att cca gtg	720
Glu Pro His Gln Lys Lys Gln Thr Arg Ala Ser Val Val Ile Pro Val	
210 215 220 225	
act ggg gat agt gaa ggt gct acg gtg cag ctg act cca tat ttt cct	768
Thr Gly Asp Ser Glu Gly Ala Thr Val Gln Leu Thr Pro Tyr Phe Pro	
230 235 240	
act tgt ggc agc gac tgc atc cga cat aaa gga aca gtt gtg ctc tgc	816
Thr Cys Gly Ser Asp Cys Ile Arg His Lys Gly Thr Val Val Leu Cys	
245 250 255	
cca caa aca ggc gtc cct ttc cct ctg gat aac aac aaa agc aag ccg	864
Pro Gln Thr Gly Val Pro Phe Pro Leu Asp Asn Asn Lys Ser Lys Pro	
260 265 270	
gga ggc tgg ctg cct ctc ctc ctg ctg tct ctg ctg gtg gcc aca tgg	912
Gly Gly Trp Leu Pro Leu Leu Leu Leu Ser Leu Leu Val Ala Thr Trp	
275 280 285	
gtg ctg gtg gca ggg atc tat cta atg tgg agg cac gaa agg atc aag	960
Val Leu Val Ala Gly Ile Tyr Leu Met Trp Arg His Glu Arg Ile Lys	
290 295 300 305	
aag act tcc ttt tct acc acc aca cta ctg ccc ccc att aag gtt ctt	1008
Lys Thr Ser Phe Ser Thr Thr Thr Leu Leu Pro Pro Ile Lys Val Leu	
310 315 320	
gtg gtt tac cca tct gaa ata tgt ttc cat cac aca att tgt tac ttc	1056
Val Val Tyr Pro Ser Glu Ile Cys Phe His His Thr Ile Cys Tyr Phe	
325 330 335	
act gaa ttt ctt caa aac cat tgc aga agt gag gtc atc ctt gaa aag	1104
Thr Glu Phe Leu Gln Asn His Cys Arg Ser Glu Val Ile Leu Glu Lys	
340 345 350	
tgg cag aaa aag aaa ata gca gag atg ggt cca gtg cag tgg ctt gcc	1152
Trp Gln Lys Lys Lys Ile Ala Glu Met Gly Pro Val Gln Trp Leu Ala	
355 360 365	
act caa aag aag gca gca gac aaa gtc gtc ttc ctt ctt tcc aat gac	1200
Thr Gln Lys Lys Ala Ala Asp Lys Val Val Phe Leu Leu Ser Asn Asp	
370 375 380 385	
gtc aac agt gtg tgc gat ggt acc tgt ggc aag agc gag ggc agt ccc	1248
Val Asn Ser Val Cys Asp Gly Thr Cys Gly Lys Ser Glu Gly Ser Pro	
390 395 400	
agt gag aac tct caa gac ctc ttc ccc ctt gcc ttt aac ctt ttc tgc	1296
Ser Glu Asn Ser Gln Asp Leu Phe Pro Leu Ala Phe Asn Leu Phe Cys	
405 410 415	
agt gat cta aga agc cag att cat ctg cac aaa tac gtg gtg gtc tac	1344
Ser Asp Leu Arg Ser Gln Ile His Leu His Lys Tyr Val Val Val Tyr	
420 425 430	



ttt aga gag att gat aca aaa gac gat tac aat gct ctc agt gtc tgc 1392  
 Phe Arg Glu Ile Asp Thr Lys Asp Asp Tyr Asn Ala Leu Ser Val Cys  
 435 440 445

ccc aag tac cac ctc atg aag gat gcc act gct ttc tgt gca gaa ctt 1440  
 Pro Lys Tyr His Leu Met Lys Asp Ala Thr Ala Phe Cys Ala Glu Leu  
 450 455 460 465

ctc cat gtc aag cag cag gtg tca gca gga aaa aga tca caa gcc tgc 1488  
 Leu His Val Lys Gln Gln Val Ser Ala Gly Lys Arg Ser Gln Ala Cys  
 470 475 480

cac gat ggc tgc tgc tcc ttg tagccacccc atgagaagca agagacctta 1539  
 His Asp Gly Cys Cys Ser Leu  
 485

aaggcttcct atcccaccaa ttacagggaa aaaacgtgtg atgatcctga agcttactat 1599  
 gcagcctaca aacagcctta gtaattaaaa cattttatac caataaaatt ttcaaattatt 1659  
 gctaactaat gtagcattaa ctaacgattg gaaactacat ttacaacttc aaagctgttt 1719  
 tatacataga aatcaattac agctttaatt gaaaactgta accattttga taatgcaaca 1779  
 ataaagcatc ttcagcc 1796

<210> 2  
 <211> 502  
 <212> PRT  
 <213> Unknown

<400> 2  
 Met Ser Leu Val Leu Leu Ser Leu Ala Ala Leu Cys Arg Ser Ala Val  
 -10 -5 -1 1

Pro Arg Glu Pro Thr Val Gln Cys Gly Ser Glu Thr Gly Pro Ser Pro  
 5 10 15

Glu Trp Met Leu Gln His Asp Leu Ile Pro Gly Asp Leu Arg Asp Leu  
 20 25 30

Arg Val Glu Pro Val Thr Thr Ser Val Ala Thr Gly Asp Tyr Ser Ile  
 35 40 45 50

Leu Met Asn Val Ser Trp Val Leu Arg Ala Asp Ala Ser Ile Arg Leu  
 55 60 65

Leu Lys Ala Thr Lys Ile Cys Val Thr Gly Lys Ser Asn Phe Gln Ser  
 70 75 80

Tyr Ser Cys Val Arg Cys Asn Tyr Thr Glu Ala Phe Gln Thr Gln Thr  
 85 90 95

Arg Pro Ser Gly Gly Lys Trp Thr Phe Ser Tyr Ile Gly Phe Pro Val  
 100 105 110

Glu Leu Asn Thr Val Tyr Phe Ile Gly Ala His Asn Ile Pro Asn Ala  
 115 120 125 130

Asn Met Asn Glu Asp Gly Pro Ser Met Ser Val Asn Phe Thr Ser Pro  
 135 140 145  
 Gly Cys Leu Asp His Ile Met Lys Tyr Lys Lys Lys Cys Val Lys Ala  
 150 155 160  
 Gly Ser Leu Trp Asp Pro Asn Ile Thr Ala Cys Lys Lys Asn Glu Glu  
 165 170 175  
 Thr Val Glu Val Asn Phe Thr Thr Thr Pro Leu Gly Asn Arg Tyr Met  
 180 185 190  
 Ala Leu Ile Gln His Ser Thr Ile Ile Gly Phe Ser Gln Val Phe Glu  
 195 200 205 210  
 Pro His Gln Lys Lys Gln Thr Arg Ala Ser Val Val Ile Pro Val Thr  
 215 220 225  
 Gly Asp Ser Glu Gly Ala Thr Val Gln Leu Thr Pro Tyr Phe Pro Thr  
 230 235 240  
 Cys Gly Ser Asp Cys Ile Arg His Lys Gly Thr Val Val Leu Cys Pro  
 245 250 255  
 Gln Thr Gly Val Pro Phe Pro Leu Asp Asn Asn Lys Ser Lys Pro Gly  
 260 265 270  
 Gly Trp Leu Pro Leu Leu Leu Leu Ser Leu Leu Val Ala Thr Trp Val  
 275 280 285 290  
 Leu Val Ala Gly Ile Tyr Leu Met Trp Arg His Glu Arg Ile Lys Lys  
 295 300 305  
 Thr Ser Phe Ser Thr Thr Thr Leu Leu Pro Pro Ile Lys Val Leu Val  
 310 315 320  
 Val Tyr Pro Ser Glu Ile Cys Phe His His Thr Ile Cys Tyr Phe Thr  
 325 330 335  
 Glu Phe Leu Gln Asn His Cys Arg Ser Glu Val Ile Leu Glu Lys Trp  
 340 345 350  
 Gln Lys Lys Lys Ile Ala Glu Met Gly Pro Val Gln Trp Leu Ala Thr  
 355 360 365 370  
 Gln Lys Lys Ala Ala Asp Lys Val Val Phe Leu Leu Ser Asn Asp Val  
 375 380 385  
 Asn Ser Val Cys Asp Gly Thr Cys Gly Lys Ser Glu Gly Ser Pro Ser  
 390 395 400  
 Glu Asn Ser Gln Asp Leu Phe Pro Leu Ala Phe Asn Leu Phe Cys Ser  
 405 410 415  
 Asp Leu Arg Ser Gln Ile His Leu His Lys Tyr Val Val Val Tyr Phe  
 420 425 430  
 Arg Glu Ile Asp Thr Lys Asp Asp Tyr Asn Ala Leu Ser Val Cys Pro  
 435 440 445 450

Lys Tyr His Leu Met Lys Asp Ala Thr Ala Phe Cys Ala Glu Leu Leu  
                   455                  460                  465

His Val Lys Gln Gln Val Ser Ala Gly Lys Arg Ser Gln Ala Cys His  
                   470                  475                  480

Asp Gly Cys Cys Ser Leu  
                   485

<210> 3  
 <211> 1506  
 <212> DNA  
 <213> reverse translation

<220>  
 <221> misc\_feature  
 <222> (1)..(1506)  
 <223> n may be a, c, g, or t

<400> 3  
 atgwsnytnng tnytnytnws nytnngcngcn ytntgymgnw sngcngtncc nmngngarccn 60  
 acngtncart gyggwnwsnga racnggnccn wsnccngart ggatgytnca rcaygayytn 120  
 athccnggng ayytnmngnga yytnmngntn garccngtna cnacnwsngt ngcnacnggn 180  
 gaytaywsna thytnatgaa ygtnwsntgg gtnytnmngng cngaygcnws nathmngnytn 240  
 ytnaargcna cnaarathtg ygtnacnggn aarwsnaayt tycarwsnta ywsntgygtn 300  
 mgntgyaayt ayacngargc nttycaracn caracnmgnc cnwsngggngg naartggacn 360  
 ttywsntaya thggnttycc ngtngarytn aayacngtnt ayttyathgg ngcncayaay 420  
 athccnaayg cnaayatgaa ygargayggn ccnwsnatgw sngtnaaytt yacnwsnccn 480  
 ggntgyytnng aycayathat gaartayaar aaraartgyg tnaargcngg nwsnytnntgg 540  
 gayccnaaya thacngcntg yaaraaraay gargaracng tngargtnaa yttyacnacn 600  
 acncnytnng gnaaymgnta yatggcnytn athcarcayw snacnathat hggnttywsn 660  
 cargtnttyg arcncayca raaraarcac acnmngncnw sngtnngtnat hccngtnacn 720  
 ggngaywsng arggngcnac ngtnrcarytn acncntayt tyccnacntg yggwnwsngay 780  
 tgyathmgnc ayaarggnac ngtnngtnytn tgyccncara cnggngtncc nttyccnytn 840  
 gayaayaaya arwsnaarcc nggnggntgg ytnccnytny tnytnytnws nytnytnngtn 900  
 gcnacntggg tnytnngtngc nggnathtay ytnatgtggm gncaygarmg nathaaraar 960  
 acnwsnttyw snacnacnac nytnytnccn ccnathaarg tnytnngtngt ntayccnwsn 1020  
 garathtgyt tycaycayac nathtgytay ttyacngart tyytncaraa ycaytgymgn 1080  
 wsgargtna thytngaraa rtggcaraar aaraarathg cngaratggg nccngtnacar 1140

tgggtngcna cncaraaraa rgcngcngay aargtngtnt tyytnytnws naaygaygt 1200  
 aaywsngtnt gygayggnac ntgyggnaar wsngarggnw snccnwsnga raaywsncar 1260  
 gayytnttyc cnytnngcmtt yaayytntty tgywsngayy tnmgnwsnca rathcayytn 1320  
 cayaartayg tngtngtnta yttymngar athgayacna argaygayta yaaygcnytn 1380  
 wsngtntgyg cnaartayca yytnatgaar gaygcnacng cnttytgygc ngarytnytn 1440  
 caygtnaarc arcargtnws ngcnggnaar mgnwsncarg cntgycayga yggntgytgy 1500  
 wsnytn 1506

<210> 4  
 <211> 637  
 <212> DNA  
 <213> Unknown

<220>  
 <223> Description of Unknown Organism:rodent; surmised  
 Mus musculus .

<220>  
 <221> CDS  
 <222> (1)..(210)

<400> 4  
 gat ttc agc agc cag acg cat ctg cac aaa tac ctg gag gtc tat ctt 48  
 Asp Phe Ser Ser Gln Thr His Leu His Lys Tyr Leu Glu Val Tyr Leu  
 1 5 10 15  
 ggg gga gca gac etc aaa ggc gac tat aat gcc ctg agt gtc tgc ccc 96  
 Gly Gly Ala Asp Leu Lys Gly Asp Tyr Asn Ala Leu Ser Val Cys Pro  
 20 25 30  
 caa tat cat ctc atg aag gac gcc aca gct ttc cac aca gaa ctt ctc 144  
 Gln Tyr His Leu Met Lys Asp Ala Thr Ala Phe His Thr Glu Leu Leu  
 35 40 45  
 aag gct acg cag agc atg tca gtg aag aaa cgc tca caa gcc tgc cat 192  
 Lys Ala Thr Gln Ser Met Ser Val Lys Lys Arg Ser Gln Ala Cys His  
 50 55 60  
 gat agc tgt tca ccc ttg tagtccaccc gggggaatag agactctgaa 240  
 Asp Ser Cys Ser Pro Leu  
 65 70  
 gccttcctac tctcccttcc agtgacaaat gctgtgtgac gactctgaaa tgtgtgggag 300  
 aggctgtgtg gaggtagtgc tatgtacaaa cttgctttaa aactggagtt tgcaaagtca 360  
 acctgagcat acacgcctga ggctagtcac tggctggatt tatgaagaca acacagttac 420  
 agacaataat gagtgggacc tacatttggg atatacccaa agctgggtaa tgattatcac 480  
 tgagaaccac gcactctggc catgaggtaa tacggcactt ccctgtcagg ctgtctgtca 540  
 ggttgggtct gtcttgact gcccatgtct tatgtgtcac gtagaccgtt ttgtaacatt 600

ttaatctgtt aatgaataat ccgtttggga ggctctc

637

<210> 5  
 <211> 70  
 <212> PRT  
 <213> Unknown

<400> 5  
 Asp Phe Ser Ser Gln Thr His Leu His Lys Tyr Leu Glu Val Tyr Leu  
     1                    5                    10                    15  
 Gly Gly Ala Asp Leu Lys Gly Asp Tyr Asn Ala Leu Ser Val Cys Pro  
                     20                    25                    30  
 Gln Tyr His Leu Met Lys Asp Ala Thr Ala Phe His Thr Glu Leu Leu  
                     35                    40                    45  
 Lys Ala Thr Gln Ser Met Ser Val Lys Lys Arg Ser Gln Ala Cys His  
                     50                    55                    60  
 Asp Ser Cys Ser Pro Leu  
     65                    70

<210> 6  
 <211> 210  
 <212> DNA  
 <213> reverse translation

<220>  
 <221> misc\_feature  
 <222> (1)..(210)  
 <223> n may be a, c, g, or t

<400> 6  
 gayttywsnw sncaracnca yytncayaar tayytngarg tntayytngg nggngcngay 60  
 ytnaarggng aytayaaygc nytnwsngtn tgyccncart aycayytnat gaargaygcn 120  
 acngcnttyc ayacngaryt nytnaargcn acncarwsna tgwsngtnaa raarmgnwsn 180  
 cargcntgyc aygaywsntg ywsnccnytn 210

<210> 7  
 <211> 2308  
 <212> DNA  
 <213> Unknown

<220>  
 <223> Description of Unknown Organism:primate; surmised  
       Homo sapiens

<220>  
 <221> CDS  
 <222> (181)..(2289)

<220>

<221> mat\_peptide  
 <222> (241)..(2289)

<220>  
 <221> misc\_feature  
 <222> (664)  
 <223> Xaa translation depends on genetic code

<400> 7

```

gagtcaggac tcccaggaca gagagtgcac aaactaccca gcacagcccc ctccgcccc 60
tctggaggct gaagagggat tccagccct gccacccaca gacacgggct gactgggggtg 120
tctgcccccc ttggggggcan ccacagggcc tcaggcctgg gtgccacctg gcactagaag 180

atg cct gtg ccc tgg ttc ttg ctg tcc ttg gca ctg ggc cga agc cag 228
Met Pro Val Pro Trp Phe Leu Leu Ser Leu Ala Leu Gly Arg Ser Gln
-20 -15 -10 -5

tgg atc ctt tct ctg gag agg ctt gtg ggg cct cag gac gct acc cac 276
Trp Ile Leu Ser Leu Glu Arg Leu Val Gly Pro Gln Asp Ala Thr His
-1 1 5 10

tgc tct ccg ggc ctc tcc tgc cgc ctc tgg gac agt gac ata ctc tgc 324
Cys Ser Pro Gly Leu Ser Cys Arg Leu Trp Asp Ser Asp Ile Leu Cys
15 20 25

ctg cct ggg gac atc gtg cct gct ccg ggc ccc gtg ctg gcg cct acg 372
Leu Pro Gly Asp Ile Val Pro Ala Pro Gly Pro Val Leu Ala Pro Thr
30 35 40

cac ctg cag aca gag ctg gtg ctg agg tgc cag aag gag acc gac tgt 420
His Leu Gln Thr Glu Leu Val Leu Arg Cys Gln Lys Glu Thr Asp Cys
45 50 55 60

gac ctc tgt ctg cgt gtg gct gtc cac ttg gcc gtg cat ggg cac tgg 468
Asp Leu Cys Leu Arg Val Ala Val His Leu Ala Val His Gly His Trp
65 70 75

gaa gag cct gaa gat gag gaa aag ttt gga gga gca gct gac tta ggg 516
Glu Glu Pro Glu Asp Glu Glu Lys Phe Gly Gly Ala Ala Asp Leu Gly
80 85 90

gtg gag gag cct agg aat gcc tct ctc cag gcc caa gtc gtg ctc tcc 564
Val Glu Glu Pro Arg Asn Ala Ser Leu Gln Ala Gln Val Val Leu Ser
95 100 105

ttc cag gcc tac cct act gcc cgc tgc gtc ctg ctg gag gtg caa gtg 612
Phe Gln Ala Tyr Pro Thr Ala Arg Cys Val Leu Leu Glu Val Gln Val
110 115 120

cct gct gcc ctt gtg cag ttt ggt cag tct gtg ggc tct gtg gta tat 660
Pro Ala Ala Leu Val Gln Phe Gly Gln Ser Val Gly Ser Val Val Tyr
125 130 135 140

gac tgc ttc gag gct gcc cta ggg agt gag gta cga atc tgg tcc tat 708
Asp Cys Phe Glu Ala Ala Leu Gly Ser Glu Val Arg Ile Trp Ser Tyr
145 150 155

act cag ccc agg tac gag aag gaa ctc aac cac aca cag cag ctg cct 756

```

Thr	Gln	Pro	Arg	Tyr	Glu	Lys	Glu	Leu	Asn	His	Thr	Gln	Gln	Leu	Pro		
			160					165					170				
gac	tgc	agg	ggg	ctc	gaa	gtc	tgg	aac	agc	atc	ccg	agc	tgc	tgg	gcc	804	
Asp	Cys	Arg	Gly	Leu	Glu	Val	Trp	Asn	Ser	Ile	Pro	Ser	Cys	Trp	Ala		
		175					180					185					
ctg	ccc	tgg	ctc	aac	gtg	tca	gca	gat	ggc	gac	aac	gtg	cat	ctg	gtt	852	
Leu	Pro	Trp	Leu	Asn	Val	Ser	Ala	Asp	Gly	Asp	Asn	Val	His	Leu	Val		
	190					195					200						
ctg	aat	gtc	tct	gag	gag	cag	cac	ttc	ggc	ctc	tcc	ctg	tac	tgg	aat	900	
Leu	Asn	Val	Ser	Glu	Glu	Gln	His	Phe	Gly	Leu	Ser	Leu	Tyr	Trp	Asn		
205				210					215					220			
cag	gtc	cag	ggc	ccc	cca	aaa	ccc	cgg	tgg	cac	aaa	aac	ctg	act	gga	948	
Gln	Val	Gln	Gly	Pro	Pro	Lys	Pro	Arg	Trp	His	Lys	Asn	Leu	Thr	Gly		
			225					230					235				
ccg	cag	atc	att	acc	ttg	aac	cac	aca	gac	ctg	gtt	ccc	tgc	ctc	tgt	996	
Pro	Gln	Ile	Ile	Thr	Leu	Asn	His	Thr	Asp	Leu	Val	Pro	Cys	Leu	Cys		
		240					245					250					
att	cag	gtg	tgg	cct	ctg	gaa	cct	gac	tcc	gtt	agg	acg	aac	atc	tgc	1044	
Ile	Gln	Val	Trp	Pro	Leu	Glu	Pro	Asp	Ser	Val	Arg	Thr	Asn	Ile	Cys		
	255					260					265						
ccc	ttc	agg	gag	gac	ccc	cgc	gca	cac	cag	aac	ctc	tgg	caa	gcc	gcc	1092	
Pro	Phe	Arg	Glu	Asp	Pro	Arg	Ala	His	Gln	Asn	Leu	Trp	Gln	Ala	Ala		
	270					275					280						
cga	ctg	cga	ctg	ctg	acc	ctg	cag	agc	tgg	ctg	ctg	gac	gca	ccg	tgc	1140	
Arg	Leu	Arg	Leu	Leu	Thr	Leu	Gln	Ser	Trp	Leu	Leu	Asp	Ala	Pro	Cys		
285					290				295					300			
tcg	ctg	ccc	gca	gaa	gcg	gca	ctg	tgc	tgg	cgg	gct	ccg	ggc	ggg	gac	1188	
Ser	Leu	Pro	Ala	Glu	Ala	Ala	Leu	Cys	Trp	Arg	Ala	Pro	Gly	Gly	Asp		
		305					310					315					
ccc	tgc	cag	cca	ctg	gtc	cca	ccg	ctt	tcc	tgg	gag	aat	gtc	act	gtg	1236	
Pro	Cys	Gln	Pro	Leu	Val	Pro	Pro	Leu	Ser	Trp	Glu	Asn	Val	Thr	Val		
		320					325					330					
gac	gtg	aac	agc	tcg	gag	aag	ctg	cag	ctg	cag	gag	tgc	ttg	tgg	gct	1284	
Asp	Val	Asn	Ser	Ser	Glu	Lys	Leu	Gln	Leu	Gln	Glu	Cys	Leu	Trp	Ala		
		335				340					345						
gac	tcc	ctg	ggg	cct	ctc	aaa	gac	gat	gtg	cta	ctg	ttg	gag	aca	cga	1332	
Asp	Ser	Leu	Gly	Pro	Leu	Lys	Asp	Asp	Val	Leu	Leu	Leu	Glu	Thr	Arg		
	350					355				360							
ggc	ccc	cag	gac	aac	aga	tcc	ctc	tgt	gcc	ttg	gaa	ccc	agt	ggc	tgt	1380	
Gly	Pro	Gln	Asp	Asn	Arg	Ser	Leu	Cys	Ala	Leu	Glu	Pro	Ser	Gly	Cys		
365				370				375						380			
act	tca	cta	ccc	agc	aaa	gcc	tcc	acg	agg	gca	gct	cgc	ctt	gga	gag	1428	
Thr	Ser	Leu	Pro	Ser	Lys	Ala	Ser	Thr	Arg	Ala	Ala	Arg	Leu	Gly	Glu		
			385					390					395				
tac	tta	cta	caa	gac	ctg	cag	tca	ggc	cag	tgt	ctg	cag	cta	tgg	gac	1476	

Tyr	Leu	Leu	Gln	Asp	Leu	Gln	Ser	Gly	Gln	Cys	Leu	Gln	Leu	Trp	Asp		
			400					405					410				
gat	gac	ttg	gga	gcg	cta	tgg	gcc	tgc	ccc	atg	gac	aaa	tac	atc	cac	1524	
Asp	Asp	Leu	Gly	Ala	Leu	Trp	Ala	Cys	Pro	Met	Asp	Lys	Tyr	Ile	His		
		415					420					425					
aag	cgc	tgg	gcc	ctc	gtg	tgg	ctg	gcc	tgc	cta	ctc	ttt	gcc	gct	gcg	1572	
Lys	Arg	Trp	Ala	Leu	Val	Trp	Leu	Ala	Cys	Leu	Leu	Phe	Ala	Ala	Ala		
	430					435					440						
ctt	tcc	ctc	atc	ctc	ctt	ctc	aaa	aag	gat	cac	gcg	aaa	ggg	tgg	ctg	1620	
Leu	Ser	Leu	Ile	Leu	Leu	Leu	Lys	Lys	Asp	His	Ala	Lys	Gly	Trp	Leu		
445					450					455					460		
agg	ctc	ttg	aaa	cag	gac	gtc	cgc	tgc	ggg	gcg	gcc	gcc	agg	ggc	cgc	1668	
Arg	Leu	Leu	Lys	Gln	Asp	Val	Arg	Ser	Gly	Ala	Ala	Ala	Arg	Gly	Arg		
			465					470					475				
gcg	gct	ctg	ctc	ctc	tac	tca	gcc	gat	gac	tgc	ggg	ttc	gag	cgc	ctg	1716	
Ala	Ala	Leu	Leu	Leu	Tyr	Ser	Ala	Asp	Asp	Ser	Gly	Phe	Glu	Arg	Leu		
		480						485					490				
gtg	ggc	gcc	ctg	gcg	tgc	gcc	ctg	tgc	cag	ctg	ccg	ctg	cgc	gtg	gcc	1764	
Val	Gly	Ala	Leu	Ala	Ser	Ala	Leu	Cys	Gln	Leu	Pro	Leu	Arg	Val	Ala		
	495					500					505						
gta	gac	ctg	tgg	agc	cgt	cgt	gaa	ctg	agc	gcg	cag	ggg	ccc	gtg	gct	1812	
Val	Asp	Leu	Trp	Ser	Arg	Arg	Glu	Leu	Ser	Ala	Gln	Gly	Pro	Val	Ala		
	510					515					520						
tgg	ttt	cac	gcg	cag	cgg	cgc	cag	acc	ctg	cag	gag	ggc	ggc	gtg	gtg	1860	
Trp	Phe	His	Ala	Gln	Arg	Arg	Gln	Thr	Leu	Gln	Glu	Gly	Gly	Val	Val		
525				530					535					540			
gtc	ttg	ctc	ttc	tct	ccc	ggg	gcg	gtg	gcg	ctg	tgc	agc	gag	tgg	cta	1908	
Val	Leu	Leu	Phe	Ser	Pro	Gly	Ala	Val	Ala	Leu	Cys	Ser	Glu	Trp	Leu		
			545					550					555				
cag	gat	ggg	gtg	tcc	ggg	ccc	ggg	gcg	cac	ggc	ccg	cac	gac	gcc	ttc	1956	
Gln	Asp	Gly	Val	Ser	Gly	Pro	Gly	Ala	His	Gly	Pro	His	Asp	Ala	Phe		
		560					565					570					
cgc	gcc	tgc	ctc	agc	tgc	gtg	ctg	ccc	gac	ttc	ttg	cag	ggc	cgg	gcg	2004	
Arg	Ala	Ser	Leu	Ser	Cys	Val	Leu	Pro	Asp	Phe	Leu	Gln	Gly	Arg	Ala		
		575				580						585					
ccc	ggc	agc	tac	gtg	ggg	gcc	tgc	ttc	gac	agg	ctg	ctc	cac	ccg	gac	2052	
Pro	Gly	Ser	Tyr	Val	Gly	Ala	Cys	Phe	Asp	Arg	Leu	Leu	His	Pro	Asp		
	590					595					600						
gcc	gta	ccc	gcc	ctt	ttc	cgc	acc	gtg	ccc	gtc	ttc	aca	ctg	ccc	tcc	2100	
Ala	Val	Pro	Ala	Leu	Phe	Arg	Thr	Val	Pro	Val	Phe	Thr	Leu	Pro	Ser		
	605				610					615					620		
caa	ctg	cca	gac	ttc	ctg	ggg	gcc	ctg	cag	cag	cct	cgc	gcc	ccg	cgt	2148	
Gln	Leu	Pro	Asp	Phe	Leu	Gly	Ala	Leu	Gln	Gln	Pro	Arg	Ala	Pro	Arg		
			625					630					635				
tcc	ggg	cgg	ctc	caa	gag	aga	gcg	gag	caa	gtg	tcc	cgg	gcc	ctt	cag	2196	



Ser Gly Arg Leu Gln Glu Arg Ala Glu Gln Val Ser Arg Ala Leu Gln  
                   640                  645                  650  
 cca gcc ctg gat agc tac ttc cat ccc ccg ggg acn tcc gcg ccg gga 2244  
 Pro Ala Leu Asp Ser Tyr Phe His Pro Pro Gly Xaa Ser Ala Pro Gly  
                   655                  660                  665  
 cgc ggg gtg gga cca ggg gcg gga cct ggg gcg ggg gac ggg act 2289  
 Arg Gly Val Gly Pro Gly Ala Gly Pro Gly Ala Gly Asp Gly Thr  
                   670                  675                  680  
 taaataaagg cagacgctg 2308  
  
 <210> 8  
 <211> 703  
 <212> PRT  
 <213> Unknown  
  
 <400> 8  
 Met Pro Val Pro Trp Phe Leu Leu Ser Leu Ala Leu Gly Arg Ser Gln  
 -20                  -15                  -10                  -5  
 Trp Ile Leu Ser Leu Glu Arg Leu Val Gly Pro Gln Asp Ala Thr His  
                   -1 1                  5                  10  
 Cys Ser Pro Gly Leu Ser Cys Arg Leu Trp Asp Ser Asp Ile Leu Cys  
                   15                  20                  25  
 Leu Pro Gly Asp Ile Val Pro Ala Pro Gly Pro Val Leu Ala Pro Thr  
                   30                  35                  40  
 His Leu Gln Thr Glu Leu Val Leu Arg Cys Gln Lys Glu Thr Asp Cys  
                   45                  50                  55                  60  
 Asp Leu Cys Leu Arg Val Ala Val His Leu Ala Val His Gly His Trp  
                   65                  70                  75  
 Glu Glu Pro Glu Asp Glu Glu Lys Phe Gly Gly Ala Ala Asp Leu Gly  
                   80                  85                  90  
 Val Glu Glu Pro Arg Asn Ala Ser Leu Gln Ala Gln Val Val Leu Ser  
                   95                  100                  105  
 Phe Gln Ala Tyr Pro Thr Ala Arg Cys Val Leu Leu Glu Val Gln Val  
                   110                  115                  120  
 Pro Ala Ala Leu Val Gln Phe Gly Gln Ser Val Gly Ser Val Val Tyr  
                   125                  130                  135                  140  
 Asp Cys Phe Glu Ala Ala Leu Gly Ser Glu Val Arg Ile Trp Ser Tyr  
                   145                  150                  155  
 Thr Gln Pro Arg Tyr Glu Lys Glu Leu Asn His Thr Gln Gln Leu Pro  
                   160                  165                  170  
 Asp Cys Arg Gly Leu Glu Val Trp Asn Ser Ile Pro Ser Cys Trp Ala  
                   175                  180                  185  
 Leu Pro Trp Leu Asn Val Ser Ala Asp Gly Asp Asn Val His Leu Val

190	195	200
Leu Asn Val Ser Glu Glu Gln His Phe Gly Leu Ser Leu Tyr Trp Asn 205 210 215 220		
Gln Val Gln Gly Pro Pro Lys Pro Arg Trp His Lys Asn Leu Thr Gly 225 230 235		
Pro Gln Ile Ile Thr Leu Asn His Thr Asp Leu Val Pro Cys Leu Cys 240 245 250		
Ile Gln Val Trp Pro Leu Glu Pro Asp Ser Val Arg Thr Asn Ile Cys 255 260 265		
Pro Phe Arg Glu Asp Pro Arg Ala His Gln Asn Leu Trp Gln Ala Ala 270 275 280		
Arg Leu Arg Leu Leu Thr Leu Gln Ser Trp Leu Leu Asp Ala Pro Cys 285 290 295 300		
Ser Leu Pro Ala Glu Ala Ala Leu Cys Trp Arg Ala Pro Gly Gly Asp 305 310 315		
Pro Cys Gln Pro Leu Val Pro Pro Leu Ser Trp Glu Asn Val Thr Val 320 325 330		
Asp Val Asn Ser Ser Glu Lys Leu Gln Leu Gln Glu Cys Leu Trp Ala 335 340 345		
Asp Ser Leu Gly Pro Leu Lys Asp Asp Val Leu Leu Leu Glu Thr Arg 350 355 360		
Gly Pro Gln Asp Asn Arg Ser Leu Cys Ala Leu Glu Pro Ser Gly Cys 365 370 375 380		
Thr Ser Leu Pro Ser Lys Ala Ser Thr Arg Ala Ala Arg Leu Gly Glu 385 390 395		
Tyr Leu Leu Gln Asp Leu Gln Ser Gly Gln Cys Leu Gln Leu Trp Asp 400 405 410		
Asp Asp Leu Gly Ala Leu Trp Ala Cys Pro Met Asp Lys Tyr Ile His 415 420 425		
Lys Arg Trp Ala Leu Val Trp Leu Ala Cys Leu Leu Phe Ala Ala Ala 430 435 440		
Leu Ser Leu Ile Leu Leu Leu Lys Lys Asp His Ala Lys Gly Trp Leu 445 450 455 460		
Arg Leu Leu Lys Gln Asp Val Arg Ser Gly Ala Ala Ala Arg Gly Arg 465 470 475		
Ala Ala Leu Leu Leu Tyr Ser Ala Asp Asp Ser Gly Phe Glu Arg Leu 480 485 490		
Val Gly Ala Leu Ala Ser Ala Leu Cys Gln Leu Pro Leu Arg Val Ala 495 500 505		
Val Asp Leu Trp Ser Arg Arg Glu Leu Ser Ala Gln Gly Pro Val Ala		

510	515	520
Trp Phe His Ala Gln Arg Arg Gln Thr Leu Gln Glu Gly Gly Val Val		
525	530	535 540
Val Leu Leu Phe Ser Pro Gly Ala Val Ala Leu Cys Ser Glu Trp Leu		
	545	550 555
Gln Asp Gly Val Ser Gly Pro Gly Ala His Gly Pro His Asp Ala Phe		
	560	565 570
Arg Ala Ser Leu Ser Cys Val Leu Pro Asp Phe Leu Gln Gly Arg Ala		
	575	580 585
Pro Gly Ser Tyr Val Gly Ala Cys Phe Asp Arg Leu Leu His Pro Asp		
	590	595 600
Ala Val Pro Ala Leu Phe Arg Thr Val Pro Val Phe Thr Leu Pro Ser		
	605	610 615 620
Gln Leu Pro Asp Phe Leu Gly Ala Leu Gln Gln Pro Arg Ala Pro Arg		
	625	630 635
Ser Gly Arg Leu Gln Glu Arg Ala Glu Gln Val Ser Arg Ala Leu Gln		
	640	645 650
Pro Ala Leu Asp Ser Tyr Phe His Pro Pro Gly Xaa Ser Ala Pro Gly		
	655	660 665
Arg Gly Val Gly Pro Gly Ala Gly Pro Gly Ala Gly Asp Gly Thr		
	670	675 680

&lt;210&gt; 9

&lt;211&gt; 2109

&lt;212&gt; DNA

&lt;213&gt; reverse translation

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1)..(2109)

&lt;223&gt; n may be a, c, g, or t

&lt;400&gt; 9

atgccngtnc cntggttyt nytnwsnytn gcnytnggnm gnwsncartg gathytnwsn 60

ytngarmgny tngtnggncc ncargaycn acncaytgyw snccnggnytn nwsntgymgn 120

ytntgggayw sngayathyt ntgyytnccn ggngayathg tncngcncc nggnccngtn 180

ytngcncna cncayytnc racngarytn gtnytnmgnt gycaraarga racngaytgy 240

gayytntggy tnmngtngc ngtncaaytn gcngtncayg gncaytggga rgarcncgar 300

gaygargara arttyggngg ngcngcngay ytngngtng argarccnmg naaygcwnsn 360

ytncargcnc argtngtnyt nwsnttycar gcntayccna cngcnmgntg ygtnytnytn 420

gargtncarg tncngcngc nytngtncar ttyggncarw sngtnggnws ngtngtntay 480

gaytgyttyg argcngcnyt nggnwsngar gtnmgnaht ggwsntayac ncarccnmgn 540  
 taygaraarg arytnaayca yacncarcar ytncngayt gymnggnyt ngargtntgg 600  
 aaywsnathc cnwsntgytg ggcnytnccn tggytnaayg tnwsngcnga yggngayaay 660  
 gtncayytng tnytnaaygt nwsngargar carcayttyg gnytnwsnyt ntaytggaay 720  
 cargtncarg gnccnecnaa rccnmgttg cayaaraay tnacnggncc ncarathath 780  
 acnytnaayc ayacngayyt ngtnccntgy ytntgyathc argtntggcc nytngarccn 840  
 gaywsngtnm gnacnaayat htgyccntty mgngargayc cnmgngcnca ycaraaytn 900  
 tggcargcng cnmgnytnmg nytnytnacn ytncarwsnt ggytnytnga ygcncntgy 960  
 wsnytnccng cngargcngc nytnntgytg mgngcncng gnggngaycc ntgycarccn 1020  
 ytngtnccnc cnytnwsntg ggaraaygt acngtngayg tnaaywsnws ngaraarytn 1080  
 carytncarg artgyytntg ggcngaywsn ytnggncny tnaargayga ygtnytnytn 1140  
 ytngaracnm gnggncnca rgayaaymgn wsnytnntgyg cnytngarcc nwsnggntgy 1200  
 acnwsnytn cwnsnaargc nwsnacnmgn gcngcnmgny tnggngarta yytnytncar 1260  
 gayytncarw snggncartg yytnarytn tgggaygayg ayytnggngc nytnntggcn 1320  
 tgyccnatgg ayaartayat hcayaarmgn tgggnytn tntggytnge ntgyytnytn 1380  
 ttygcngcng cnytnwsnyt nathytnytn ytnaaraarg aycaygcnaa rggntggytn 1440  
 mgnytnytna arcargaygt nmgnwsnggn gcngcngcnm gnggnmgngc ngcnytnytn 1500  
 ytntaywsng cngaygayws nggnttygar mgnytngtng gngcnytnge nwsngcnytn 1560  
 tgyarytn cnytnmgnt ngcngtngay ytntggwsnm gnmngaryt nwsngcncar 1620  
 ggcncngtng cntggttyca ygcncarmgn mgncaracny tncargargg nggngtngtn 1680  
 gtnytnytn tywsncngg ngcngtngcn ytntgywsng artggytnca rgayggngtn 1740  
 wsnggncng gngcncaygg nccncaygay gcnttymng cwnsnytnws ntgygtnytn 1800  
 ccngaytty tncarggnmg ngcncnggn wsntaygtng gngcntgytt ygaymgnytn 1860  
 ytncayccng aygngtnc ngcnytnntty mgnacngtnc cngtnttyac nytnccnwsn 1920  
 carytnccng ayttyytnng ngcnytnar carccnmng cncnmgng nngnmgnytn 1980  
 cargarmng cngarcargt nwsnmngcn ytncarccng cnytngayws ntayttycay 2040  
 ccncnggna cwnsngcnc nggngmgngn gtnggncng gngcnggncc nggngcnggn 2100  
 gayggnacn 2109

<210> 10  
 <211> 2314  
 <212> DNA

<213> Unknown

<220>

<223> Description of Unknown Organism:rodent; surmised  
Mus musculus

<220>

<221> CDS

<222> (199) .. (2292)

<220>

<221> mat\_peptide

<222> (259) .. (2292)

<400> 10

```

ccaaatcgaa agcacgggag ctgatactgg gcctggagtc caggctcact ggagtgggga 60

agcatggctg gagaggaatt ctagcccttg ctctctccca gggacacggg gctgattgtc 120

agcaggggag aggggtctgc ccccccttgg gggggcagga cggggcctca ggcctgggtg 180

ctgtccggca cctggaag atg cct gtg tcc tgg ttc ctg ctg tcc ttg gca 231
      Met Pro Val Ser Trp Phe Leu Leu Ser Leu Ala
      -20                               -15               -10

ctg ggc cga aac cct gtg gtc gtc tct ctg gag aga ctg atg gag cct 279
Leu Gly Arg Asn Pro Val Val Val Ser Leu Glu Arg Leu Met Glu Pro
      -5                               -1    1               5

cag gac act gca cgc tgc tct cta ggc ctc tcc tgc cac ctc tgg gat 327
Gln Asp Thr Ala Arg Cys Ser Leu Gly Leu Ser Cys His Leu Trp Asp
      10                               15               20

ggg gac gtg ctc tgc ctg cct gga agc ctc cag tct gcc cca ggc cct 375
Gly Asp Val Leu Cys Leu Pro Gly Ser Leu Gln Ser Ala Pro Gly Pro
      25                               30               35

gtg cta gtg cct acc cgc ctg cag acg gag ctg gtg ctg agg tgt cca 423
Val Leu Val Pro Thr Arg Leu Gln Thr Glu Leu Val Leu Arg Cys Pro
      40                               45               50               55

cag aag aca gat tgc gcc ctc tgt gtc cgt gtg gtg gtc cac ttg gcc 471
Gln Lys Thr Asp Cys Ala Leu Cys Val Arg Val Val Val His Leu Ala
      60                               65               70

gtg cat ggg cac tgg gca gag cct gaa gaa gct gga aag tct gat tca 519
Val His Gly His Trp Ala Glu Pro Glu Glu Ala Gly Lys Ser Asp Ser
      75                               80               85

gaa ctc cag gag tct agg aac gcc tct ctc cag gcc cag gtg gtg ctc 567
Glu Leu Gln Glu Ser Arg Asn Ala Ser Leu Gln Ala Gln Val Val Leu
      90                               95               100

tcc ttc cag gcc tac ccc atc gcc cgc tgt gcc ctg ctg gag gtc cag 615
Ser Phe Gln Ala Tyr Pro Ile Ala Arg Cys Ala Leu Leu Glu Val Gln
      105                               110               115

gtg ccc gct gac ctg gtg cag cct ggt cag tcc gtg ggt tct gcg gta 663
Val Pro Ala Asp Leu Val Gln Pro Gly Gln Ser Val Gly Ser Ala Val
      120                               125               130               135

```

ttt gac tgt ttc gag gct agt ctt ggg gct gag gta cag atc tgg tcc	711
Phe Asp Cys Phe Glu Ala Ser Leu Gly Ala Glu Val Gln Ile Trp Ser	
140 145 150	
tac acg aag ccc agg tac cag aaa gag ctc aac ctc aca cag cag ctg	759
Tyr Thr Lys Pro Arg Tyr Gln Lys Glu Leu Asn Leu Thr Gln Gln Leu	
155 160 165	
cct gac tgc agg ggt ctt gaa gtc cgg gac agc atc cag agc tgc tgg	807
Pro Asp Cys Arg Gly Leu Glu Val Arg Asp Ser Ile Gln Ser Cys Trp	
170 175 180	
gtc ctg ccc tgg ctc aat gtg tct aca gat ggt gac aat gtc ctt ctg	855
Val Leu Pro Trp Leu Asn Val Ser Thr Asp Gly Asp Asn Val Leu Leu	
185 190 195	
aca ctg gat gtc tct gag gag cag gac ttt agc ttc tta ctg tac ctg	903
Thr Leu Asp Val Ser Glu Glu Gln Asp Phe Ser Phe Leu Leu Tyr Leu	
200 205 210 215	
cgt cca gtc ccg gat gct ctc aaa tcc ttg tgg tac aaa aac ctg act	951
Arg Pro Val Pro Asp Ala Leu Lys Ser Leu Trp Tyr Lys Asn Leu Thr	
220 225 230	
gga cct cag aac att act tta aac cac aca gac ctg gtt ccc tgc ctc	999
Gly Pro Gln Asn Ile Thr Leu Asn His Thr Asp Leu Val Pro Cys Leu	
235 240 245	
tgc att cag gtg tgg tcg cta gag cca gac tct gag agg gtc gaa ttc	1047
Cys Ile Gln Val Trp Ser Leu Glu Pro Asp Ser Glu Arg Val Glu Phe	
250 255 260	
tgc ccc ttc cgg gaa gat ccc ggt gca cac agg aac ctc tgg cac ata	1095
Cys Pro Phe Arg Glu Asp Pro Gly Ala His Arg Asn Leu Trp His Ile	
265 270 275	
gcc agg ctg cgg gta ctg tcc cca ggg gta tgg cag cta gat gcg cct	1143
Ala Arg Leu Arg Val Leu Ser Pro Gly Val Trp Gln Leu Asp Ala Pro	
280 285 290 295	
tgc tgt ctg ccg ggc aag gta aca ctg tgc tgg cag gca cca gac cag	1191
Cys Cys Leu Pro Gly Lys Val Thr Leu Cys Trp Gln Ala Pro Asp Gln	
300 305 310	
agt ccc tgc cag cca ctt gtg cca cca gtg ccc cag aag aac gcc act	1239
Ser Pro Cys Gln Pro Leu Val Pro Pro Val Pro Gln Lys Asn Ala Thr	
315 320 325	
gtg aat gag cca caa gat ttc cag ttg gtg gca ggc cac ccc aac ctc	1287
Val Asn Glu Pro Gln Asp Phe Gln Leu Val Ala Gly His Pro Asn Leu	
330 335 340	
tgt gtc cag gtg agc acc tgg gag aag gtt cag ctg caa gcg tgc ttg	1335
Cys Val Gln Val Ser Thr Trp Glu Lys Val Gln Leu Gln Ala Cys Leu	
345 350 355	
tgg gct gac tcc ttg ggg ccc ttc aag gat gat atg ctg tta gtg gag	1383
Trp Ala Asp Ser Leu Gly Pro Phe Lys Asp Asp Met Leu Leu Val Glu	
360 365 370 375	

atg aaa acc ggc ctc aac aac aca tca gtc tgt gcc ttg gaa ccc agt	1431
Met Lys Thr Gly Leu Asn Asn Thr Ser Val Cys Ala Leu Glu Pro Ser	
380 385 390	
ggc tgt aca cca ctg ccc agc atg gcc tcc acg aga gct gct cgc ctg	1479
Gly Cys Thr Pro Leu Pro Ser Met Ala Ser Thr Arg Ala Ala Arg Leu	
395 400 405	
gga gag gag ttg ctg caa gac ttc cga tca cac cag tgt atg cag ctg	1527
Gly Glu Glu Leu Leu Gln Asp Phe Arg Ser His Gln Cys Met Gln Leu	
410 415 420	
tgg aac gat gac aac atg gga tgc cta tgg gcc tgc ccc atg gac aag	1575
Trp Asn Asp Asp Asn Met Gly Ser Leu Trp Ala Cys Pro Met Asp Lys	
425 430 435	
tac atc cac agg cgc tgg gtc cta gta tgg ctg gcc tgc cta ctc ttg	1623
Tyr Ile His Arg Arg Trp Val Leu Val Trp Leu Ala Cys Leu Leu Leu	
440 445 450 455	
gct gcg gcg ctt ttc ttc ttc ctc ctt cta aaa aag gac cgc agg aaa	1671
Ala Ala Ala Leu Phe Phe Phe Leu Leu Leu Lys Lys Asp Arg Arg Lys	
460 465 470	
gcg gcc cgt ggc tcc cgc acg gcc ttg ctc ctc cac tcc gcc gac gga	1719
Ala Ala Arg Gly Ser Arg Thr Ala Leu Leu Leu His Ser Ala Asp Gly	
475 480 485	
gcg ggc tac gag cgc ctg gtg gga gca ctg gcg tcc gcg ttg agc cag	1767
Ala Gly Tyr Glu Arg Leu Val Gly Ala Leu Ala Ser Ala Leu Ser Gln	
490 495 500	
atg cca ctg cgc gtg gcc gtg gac ctg tgg agc cgc cgc gag ctg agc	1815
Met Pro Leu Arg Val Ala Val Asp Leu Trp Ser Arg Arg Glu Leu Ser	
505 510 515	
gcg cac gga gcc cta gcc tgg ttc cac cac cag cga cgc cgt atc ctg	1863
Ala His Gly Ala Leu Ala Trp Phe His His Gln Arg Arg Arg Ile Leu	
520 525 530 535	
cag gag ggt ggc gtg gta atc ctt ctc ttc tgc ccc gcg gcc gtg gcg	1911
Gln Glu Gly Gly Val Val Ile Leu Leu Phe Ser Pro Ala Ala Val Ala	
540 545 550	
cag tgt cag cag tgg ctg cag ctc cag aca gtg gag ccc ggg ccg cat	1959
Gln Cys Gln Gln Trp Leu Gln Leu Gln Thr Val Glu Pro Gly Pro His	
555 560 565	
gac gcc ctc gcc gcc tgg ctc agc tgc gtg cta ccc gat ttc ctg caa	2007
Asp Ala Leu Ala Ala Trp Leu Ser Cys Val Leu Pro Asp Phe Leu Gln	
570 575 580	
ggc cgg gcg acc ggc cgc tac gtc ggg gtc tac ttc gac ggg ctg ctg	2055
Gly Arg Ala Thr Gly Arg Tyr Val Gly Val Tyr Phe Asp Gly Leu Leu	
585 590 595	
cac cca gac tct gtg ccc tcc ccg ttc cgc gtc gcc ccg ctc ttc tcc	2103
His Pro Asp Ser Val Pro Ser Pro Phe Arg Val Ala Pro Leu Phe Ser	
600 605 610 615	

```

ctg ccc tcg cag ctg ccg gct ttc ctg gat gca ctg cag gga ggc tgc 2151
Leu Pro Ser Gln Leu Pro Ala Phe Leu Asp Ala Leu Gln Gly Gly Cys
                620                625                630

tcc act tcc gcg ggg cga ccc gcg gac cgg gtg gaa cga gtg acc cag 2199
Ser Thr Ser Ala Gly Arg Pro Ala Asp Arg Val Glu Arg Val Thr Gln
                635                640                645

gcg ctg cgg tcc gcc ctg gac agc tgt act tct agc tcg gaa gcc cca 2247
Ala Leu Arg Ser Ala Leu Asp Ser Cys Thr Ser Ser Ser Glu Ala Pro
                650                655                660

ggc tgc tgc gag gaa tgg gac ctg gga ccc tgc act aca cta gaa 2292
Gly Cys Cys Glu Glu Trp Asp Leu Gly Pro Cys Thr Thr Leu Glu
                665                670                675

taaaagccga tacagtattc ct 2314

```

```

<210> 11
<211> 698
<212> PRT
<213> Unknown

```

```

<400> 11
Met Pro Val Ser Trp Phe Leu Leu Ser Leu Ala Leu Gly Arg Asn Pro
-20                -15                -10                -5

Val Val Val Ser Leu Glu Arg Leu Met Glu Pro Gln Asp Thr Ala Arg
                -1    1                5                10

Cys Ser Leu Gly Leu Ser Cys His Leu Trp Asp Gly Asp Val Leu Cys
                15                20                25

Leu Pro Gly Ser Leu Gln Ser Ala Pro Gly Pro Val Leu Val Pro Thr
                30                35                40

Arg Leu Gln Thr Glu Leu Val Leu Arg Cys Pro Gln Lys Thr Asp Cys
                45                50                55                60

Ala Leu Cys Val Arg Val Val Val His Leu Ala Val His Gly His Trp
                65                70                75

Ala Glu Pro Glu Glu Ala Gly Lys Ser Asp Ser Glu Leu Gln Glu Ser
                80                85                90

Arg Asn Ala Ser Leu Gln Ala Gln Val Val Leu Ser Phe Gln Ala Tyr
                95                100                105

Pro Ile Ala Arg Cys Ala Leu Leu Glu Val Gln Val Pro Ala Asp Leu
                110                115                120

Val Gln Pro Gly Gln Ser Val Gly Ser Ala Val Phe Asp Cys Phe Glu
                125                130                135                140

Ala Ser Leu Gly Ala Glu Val Gln Ile Trp Ser Tyr Thr Lys Pro Arg
                145                150                155

Tyr Gln Lys Glu Leu Asn Leu Thr Gln Gln Leu Pro Asp Cys Arg Gly

```



160										165					170				
Leu	Glu	Val	Arg	Asp	Ser	Ile	Gln	Ser	Cys	Trp	Val	Leu	Pro	Trp	Leu				
		175					180					185							
Asn	Val	Ser	Thr	Asp	Gly	Asp	Asn	Val	Leu	Leu	Thr	Leu	Asp	Val	Ser				
	190					195					200								
Glu	Glu	Gln	Asp	Phe	Ser	Phe	Leu	Leu	Tyr	Leu	Arg	Pro	Val	Pro	Asp				
205					210					215					220				
Ala	Leu	Lys	Ser	Leu	Trp	Tyr	Lys	Asn	Leu	Thr	Gly	Pro	Gln	Asn	Ile				
				225				230						235					
Thr	Leu	Asn	His	Thr	Asp	Leu	Val	Pro	Cys	Leu	Cys	Ile	Gln	Val	Trp				
			240				245						250						
Ser	Leu	Glu	Pro	Asp	Ser	Glu	Arg	Val	Glu	Phe	Cys	Pro	Phe	Arg	Glu				
	255					260						265							
Asp	Pro	Gly	Ala	His	Arg	Asn	Leu	Trp	His	Ile	Ala	Arg	Leu	Arg	Val				
	270					275				280									
Leu	Ser	Pro	Gly	Val	Trp	Gln	Leu	Asp	Ala	Pro	Cys	Cys	Leu	Pro	Gly				
285				290					295						300				
Lys	Val	Thr	Leu	Cys	Trp	Gln	Ala	Pro	Asp	Gln	Ser	Pro	Cys	Gln	Pro				
			305					310						315					
Leu	Val	Pro	Pro	Val	Pro	Gln	Lys	Asn	Ala	Thr	Val	Asn	Glu	Pro	Gln				
		320					325						330						
Asp	Phe	Gln	Leu	Val	Ala	Gly	His	Pro	Asn	Leu	Cys	Val	Gln	Val	Ser				
	335					340						345							
Thr	Trp	Glu	Lys	Val	Gln	Leu	Gln	Ala	Cys	Leu	Trp	Ala	Asp	Ser	Leu				
	350				355						360								
Gly	Pro	Phe	Lys	Asp	Asp	Met	Leu	Leu	Val	Glu	Met	Lys	Thr	Gly	Leu				
365				370					375						380				
Asn	Asn	Thr	Ser	Val	Cys	Ala	Leu	Glu	Pro	Ser	Gly	Cys	Thr	Pro	Leu				
			385					390						395					
Pro	Ser	Met	Ala	Ser	Thr	Arg	Ala	Ala	Arg	Leu	Gly	Glu	Glu	Leu	Leu				
		400					405						410						
Gln	Asp	Phe	Arg	Ser	His	Gln	Cys	Met	Gln	Leu	Trp	Asn	Asp	Asp	Asn				
	415					420						425							
Met	Gly	Ser	Leu	Trp	Ala	Cys	Pro	Met	Asp	Lys	Tyr	Ile	His	Arg	Arg				
	430				435						440								
Trp	Val	Leu	Val	Trp	Leu	Ala	Cys	Leu	Leu	Leu	Ala	Ala	Ala	Leu	Phe				
445				450					455					460					
Phe	Phe	Leu	Leu	Leu	Lys	Lys	Asp	Arg	Arg	Lys	Ala	Ala	Arg	Gly	Ser				
			465				470							475					
Arg	Thr	Ala	Leu	Leu	Leu	His	Ser	Ala	Asp	Gly	Ala	Gly	Tyr	Glu	Arg				

```
<400> 12
atgccngtnw sntgggttyt nytnwsnytn gcnytnggnm gnaayccngt ngtnngtnwsn 60
ytnfarmgny tnatggarcc ncargayacn gcnmgntgyw snytnggny t nwsntgycay 120
ytntgggag y gngaygtnyt ntgyytnccn ggnwsnytn arwsngcncc nggnccngtn 180
ytngtnccna cnmgnytnca racngarytn gtnytnmgnt gyccncaraa racngaytgy 240
gcnytnctgy tnmqngtngt ngtncayytn gcngtncayg gncaytgggc ngarccngar 300
```

gargcnggna arwsngayws ngarytnear garwsnmgna aygcnwsnyt ncargcnear 360  
gtngtnytnw snttycargc ntayccnath gcnmgntgyg cnytnytnga rgtncargtn 420  
ccngcngayy tngtncarcc nggncarwsn gtnggnwsng cngtnttyga ytgyttygar 480  
gcnwsnytnng gngcngargt ncarathtgg wsntayacna arccnmgnta ycaraargar 540  
ytnaaaytna cncarcaryt nccngaytgy mgnggnytnng argtnmgnga ywsnathcar 600  
wsntgytggg tnytnccntg gytnaaygt nwsnacngayg gngayaaygt nytnytnacn 660  
ytngaygt nwsngargarca rgayttywsn ttyytnytnt ayytnmgnc ngtnccngay 720  
gcnytnaarw snytntggt yaaraaytn acnggnccnc araayathac nytnaaycay 780  
acngayytnng tnccntgyt ntgyathcar gtntggwsny tngarccnga ywsngarmgn 840  
gtngarttyt gyccnttymg ngargayccn ggngcncaym gnaayytnng gcayathgc 900  
mgnytnmgng tnytnwsncc ngngtntgg carytngayg cncntgytg yytnccnggn 960  
aargtnacny tntgytggca rgcnccngay carwsnccnt gycarccnyt ngtnccnccn 1020  
gtncncara araaygcnc ngtnaaygar ccncargayt tycarytngt ngcnggncay 1080  
ccnaayytn gygtncargt nwsnacntgg garaargtnc arytncargc ntgyytntg 1140  
gngaywsny tngnccntt yaargaygay atgytnytng tngaratgaa racnggnytn 1200  
aayaayacnw sngtntgygc nytngarccn wsnggntgya cncnytncc nwsnatggcn 1260  
wsnacnmng cngcnmgnyt ngngargar ytnytnargc aytymgnws ncaycartgy 1320  
atgcarytn ggaaygayga yaayatgggn wsnytntggt cntgyccnat ggayaartay 1380  
athcaymgnm gntgggtnyt ngtntggytn gcntgyytny tnytnngcngc ngcnytnnty 1440  
ttyttytny tnytnaaraa rgaymgnmgn aargcngcnm gnggnwsnmg nacngcnytn 1500  
ytnytncayw sngcngaygg ngcngntay garmgnytn tngngcnytn ngcnwsngcn 1560  
ytnwsncara tgcnytnmg ngtnngcngtn gayytntgw smgmngnga ryt nwsngcn 1620  
cayggngcny tngcntggt ycaycaycar mgnmngmna thytncarga rggngngtn 1680  
gtnathytny tnttywsncc ngcngcngtn gcncartgy arcartggyt ncarytnear 1740  
acngtnargc cnggnccnca ygaygnytn gcngcntggy tnwsntgyt nytnccngay 1800  
ttyytnarg gnmngcnc nggnmgntay gtngngntnt ayttygaygg nytnytnca 1860  
ccngaywsng tnccnwsncc nttymgntn gcncnytn tywsnytncc nwsncarytn 1920  
ccngcnttyt ngaygnytn ncargnggn tgywsnacw sngcnggmng nccngcngay 1980  
mgngtnargm gngtnacna rgcnytnmgn wsngcnytn aywsntgyac nwsnwsnwsn 2040  
gargcncng gntgytgyga rgartgggay ytngnccnt gyacnacnytn ngar 2094

<210> 13  
 <211> 2786  
 <212> DNA  
 <213> Unknown

<220>  
 <223> Description of Unknown Organism: primate; surmised  
 Homo sapiens

<220>  
 <221> CDS  
 <222> (70)..(2283)

<220>  
 <221> mat\_peptide  
 <222> (118)..(2283)

<220>  
 <221> misc\_feature  
 <222> (9)..(134)  
 <223> Xaa translation (9, 18, 26, 109, 120, 134) depends  
 on genetic code

<400> 13  
 cccacgcntc cgggccagca gcgggcggcc ggggcgcaga gaacggcctg gctgggcgag 60

cgcacggcc atg gcc ccg tgg ctg cag ctc tgc tcc gtc ttc ttt acg gtc 111  
 Met Ala Pro Trp Leu Gln Leu Cys Ser Val Phe Phe Thr Val  
 -15 -10 -5

aac gcc tgc ctc aac ggc tcg cag ctg gct gtn gcc gct ggc ggg tcc 159  
 Asn Ala Cys Leu Asn Gly Ser Gln Leu Ala Xaa Ala Ala Gly Gly Ser  
 -1 1 5 10

ggc cgc gcg cng ggc gcc gac acc tgt agc tgg ang gga gtg ggg cca 207  
 Gly Arg Ala Xaa Gly Ala Asp Thr Cys Ser Trp Xaa Gly Val Gly Pro  
 15 20 25 30

gcc agc aga aac agt ggg ctg tac aac atc acc ttc aaa tat gac aat 255  
 Ala Ser Arg Asn Ser Gly Leu Tyr Asn Ile Thr Phe Lys Tyr Asp Asn  
 35 40 45

tgt acc acc tac ttg aat cca gtg ggg aag cat gtg att gct gac gcc 303  
 Cys Thr Thr Tyr Leu Asn Pro Val Gly Lys His Val Ile Ala Asp Ala  
 50 55 60

cag aat atc acc atc agc cag tat gct tgc cat gac caa gtg gca gtc 351  
 Gln Asn Ile Thr Ile Ser Gln Tyr Ala Cys His Asp Gln Val Ala Val  
 65 70 75

acc att ctt tgg tcc cca ggg gcc ctc ggc atc gaa ttc ctg aaa gga 399  
 Thr Ile Leu Trp Ser Pro Gly Ala Leu Gly Ile Glu Phe Leu Lys Gly  
 80 85 90

ttt cgg gta ata ctg gag gag ctg aag tcg gag gga aga cag ngc caa 447  
 Phe Arg Val Ile Leu Glu Leu Lys Ser Glu Gly Arg Gln Xaa Gln  
 95 100 105 110

caa ctg att cta aag gat ccg aag cag ntc aac agt agc ttc aaa aga 495

Gln	Leu	Ile	Leu	Lys	Asp	Pro	Lys	Gln	Xaa	Asn	Ser	Ser	Phe	Lys	Arg		
				115					120					125			
act	gga	atg	gaa	tct	caa	cct	ttt	ctg	aat	atg	aaa	ttt	gaa	acg	gat	543	
Thr	Gly	Met	Glu	Ser	Gln	Pro	Xaa	Leu	Asn	Met	Lys	Phe	Glu	Thr	Asp		
			130					135					140				
tat	ttc	gta	agg	ttg	tcc	ttt	tcc	ttc	att	aaa	aac	gaa	agc	aat	tac	591	
Tyr	Phe	Val	Arg	Leu	Ser	Phe	Ser	Phe	Ile	Lys	Asn	Glu	Ser	Asn	Tyr		
		145						150				155					
cac	cct	ttc	ttc	ttt	aga	acc	cga	gcc	tgt	gac	ctg	ttg	tta	cag	ccg	639	
His	Pro	Phe	Phe	Phe	Arg		Arg	Ala	Cys	Asp	Leu	Leu	Leu	Gln	Pro		
		160					165				170						
gac	aat	cta	gct	tgt	aaa	ccc	ttc	tgg	aag	cct	cgg	aac	ctg	aac	atc	687	
Asp	Asn	Leu	Ala	Cys	Lys	Pro	Phe	Trp	Lys	Pro	Arg	Asn	Leu	Asn	Ile		
175					180					185					190		
agc	cag	cat	ggc	tcg	gac	atg	cag	gtg	tcc	ttc	gac	cac	gca	ccg	cac	735	
Ser	Gln	His	Gly	Ser	Asp	Met	Gln	Val	Ser	Phe	Asp	His	Ala	Pro	His		
				195					200					205			
aac	ttc	ggc	ttc	cgt	ttc	ttc	tat	ctt	cac	tac	aag	ctc	aag	cac	gaa	783	
Asn	Phe	Gly	Phe	Arg	Phe	Phe	Tyr	Leu	His	Tyr	Lys	Leu	Lys	His	Glu		
			210					215					220				
gga	cct	ttc	aag	cga	aag	acc	tgt	aag	cag	gag	caa	act	aca	gag	atg	831	
Gly	Pro	Phe	Lys	Arg	Lys	Thr	Cys	Lys	Gln	Glu	Gln	Thr	Thr	Glu	Met		
		225					230					235					
acc	agc	tgc	ctc	ctt	caa	aat	gtt	tct	cca	ggg	gat	tat	ata	att	gag	879	
Thr	Ser	Cys	Leu	Leu	Gln	Asn	Val	Ser	Pro	Gly	Asp	Tyr	Ile	Ile	Glu		
		240				245					250						
ctg	gtg	gat	gac	act	aac	aca	aca	aga	aaa	gtg	atg	cat	tat	gcc	tta	927	
Leu	Val	Asp	Asp	Thr	Asn	Thr	Thr	Arg	Lys	Val	Met	His	Tyr	Ala	Leu		
255					260					265					270		
aag	cca	gtg	cac	tcc	ccg	tgg	gcc	ggg	ccc	atc	aga	gcc	gtg	gcc	atc	975	
Lys	Pro	Val	His	Ser	Pro	Trp	Ala	Gly	Pro	Ile	Arg	Ala	Val	Ala	Ile		
				275					280					285			
aca	gtg	cca	ctg	gta	gtc	ata	tcg	gca	ttc	gcg	acg	ctc	ttc	act	gtg	1023	
Thr	Val	Pro	Leu	Val	Val	Ile	Ser	Ala	Phe	Ala	Thr	Leu	Phe	Thr	Val		
			290					295					300				
atg	tgc	cgc	aag	aag	caa	caa	gaa	aat	ata	tat	tca	cat	tta	gat	gaa	1071	
Met	Cys	Arg	Lys	Lys	Gln	Gln	Glu	Asn	Ile	Tyr	Ser	His	Leu	Asp	Glu		
		305					310					315					
gag	agc	tct	gag	tct	tcc	aca	tac	act	gca	gca	ctc	cca	aga	gag	agg	1119	
Glu	Ser	Ser	Glu	Ser	Ser	Thr	Tyr	Thr	Ala	Ala	Leu	Pro	Arg	Glu	Arg		
		320				325					330						
ctc	cgg	ccg	cgg	ccg	aag	gtc	ttt	ctc	tgc	tat	tcc	agt	aaa	gat	ggc	1167	
Leu	Arg	Pro	Arg	Pro	Lys	Val	Phe	Leu	Cys	Tyr	Ser	Ser	Lys	Asp	Gly		
335					340				345					350			
cag	aat	cac	atg	aat	gtc	gtc	cag	tgt	ttc	gcc	tac	ttc	ctc	cag	gac	1215	

Gln Asn His Met Asn Val Val Gln Cys Phe Ala Tyr Phe Leu Gln Asp	
355 360 365	
ttc tgt ggc tgt gag gtg gct ctg gac ctg tgg gaa gac ttc agc ctc	1263
Phe Cys Gly Cys Glu Val Ala Leu Asp Leu Trp Glu Asp Phe Ser Leu	
370 375 380	
tgt aga gaa ggg cag aga gaa tgg gtc atc cag aag atc cac gag tcc	1311
Cys Arg Glu Gly Gln Arg Glu Trp Val Ile Gln Lys Ile His Glu Ser	
385 390 395	
cag ttc atc att gtg gtt tgt tcc aaa ggt atg aag tac ttt gtg gac	1359
Gln Phe Ile Ile Val Val Cys Ser Lys Gly Met Lys Tyr Phe Val Asp	
400 405 410	
aag aag aac tac aaa cac aaa gga ggt ggc cga ggc tcg ggg aaa gga	1407
Lys Lys Asn Tyr Lys His Lys Gly Gly Gly Arg Gly Ser Gly Lys Gly	
415 420 425 430	
gag ctc ttc ctg gtg gcg gtg tca gcc att gcc gaa aag ctc cgc cag	1455
Glu Leu Phe Leu Val Ala Val Ser Ala Ile Ala Glu Lys Leu Arg Gln	
435 440 445	
gcc aag cag agt tgc tcc gcg gcg ctc agc aag ttt atc gcc gtc tac	1503
Ala Lys Gln Ser Ser Ser Ala Ala Leu Ser Lys Phe Ile Ala Val Tyr	
450 455 460	
ttt gat tat tcc tgc gag gga gac gtc ccc ggt atc cta gac ctg agt	1551
Phe Asp Tyr Ser Cys Glu Gly Asp Val Pro Gly Ile Leu Asp Leu Ser	
465 470 475	
acc aag tac aga ctc atg gac aat ctt cct cag ctc tgt tcc cac ctg	1599
Thr Lys Tyr Arg Leu Met Asp Asn Leu Pro Gln Leu Cys Ser His Leu	
480 485 490	
cac tcc cga gac cac ggc ctc cag gag ccg ggg cag cac acg cga cag	1647
His Ser Arg Asp His Gly Leu Gln Glu Pro Gly Gln His Thr Arg Gln	
495 500 505 510	
ggc agc aga agg aac tac ttc cgg agc aag tca ggc cgg tcc cta tac	1695
Gly Ser Arg Arg Asn Tyr Phe Arg Ser Lys Ser Gly Arg Ser Leu Tyr	
515 520 525	
gtc gcc att tgc aac atg cac cag ttt att gac gag gag ccc gac tgg	1743
Val Ala Ile Cys Asn Met His Gln Phe Ile Asp Glu Glu Pro Asp Trp	
530 535 540	
ttc gaa aag cag ttc gtt ccc ttc cat cct cct cca ctg cgc tac cgg	1791
Phe Glu Lys Gln Phe Val Pro Phe His Pro Pro Pro Leu Arg Tyr Arg	
545 550 555	
gag cca gtc ttg gag aaa ttt gat tgc ggc ttg gtt tta aat gat gtc	1839
Glu Pro Val Leu Glu Lys Phe Asp Ser Gly Leu Val Leu Asn Asp Val	
560 565 570	
atg tgc aaa cca ggg cct gag agt gac ttc tgc cta aag gta gag gcg	1887
Met Cys Lys Pro Gly Pro Glu Ser Asp Phe Cys Leu Lys Val Glu Ala	
575 580 585 590	
gct gtt ctt ggg gca acc gga cca gcc gac tcc cag cac gag agt cag	1935

Ala Val Leu Gly Ala Thr Gly Pro Ala Asp Ser Gln His Glu Ser Gln  
595 600 605

cat ggg ggc ctg gac caa gac ggg gag gcc cgg cct gcc ctt gac ggt 1983  
His Gly Gly Leu Asp Gln Asp Gly Glu Ala Arg Pro Ala Leu Asp Gly  
610 615 620

agc gcc gcc ctg caa ccc ctg ctg cac acg gtg aaa gcc ggc agc ccc 2031  
Ser Ala Ala Leu Gln Pro Leu Leu His Thr Val Lys Ala Gly Ser Pro  
625 630 635

tcg gac atg ccg cgg gac tca ggc atc tat gac tcg tct gtg ccc tca 2079  
Ser Asp Met Pro Arg Asp Ser Gly Ile Tyr Asp Ser Ser Val Pro Ser  
640 645 650

tcc gag ctg tct ctg cca ctg atg gaa gga ctc tcg acg gac cag aca 2127  
Ser Glu Leu Ser Leu Pro Leu Met Glu Gly Leu Ser Thr Asp Gln Thr  
655 660 665 670

gaa acg tct tcc ctg acg gag agc gtg tcc tcc tct tca ggc ctg ggt 2175  
Glu Thr Ser Ser Leu Thr Glu Ser Val Ser Ser Ser Ser Gly Leu Gly  
675 680 685

gag gag gaa cct cct gcc ctt cct tcc aag ctc ctc tct tct ggg tca 2223  
Glu Glu Glu Pro Pro Ala Leu Pro Ser Lys Leu Leu Ser Ser Gly Ser  
690 695 700

tgc aaa gca gat ctt ggt tgc cgc agc tac act gat gaa ctc cac gcg 2271  
Cys Lys Ala Asp Leu Gly Cys Arg Ser Tyr Thr Asp Glu Leu His Ala  
705 710 715

gtc gcc cct ttg taacaaaacg aaagagtcta agcattgccca ctttagctgc 2323  
Val Ala Pro Leu  
720

tgcctccctc tgattcccca gctcatctcc ctggttgcac ggcccacttg gagctgaggt 2383

ctcatacaag gatatttgga gtgaaatgct ggccagtact tggtctccct tgccccaacc 2443

ctttaccgga tatcttgaca aactctccaa ttttctaaaa tgatatggag ctctgaaagg 2503

catgtccata aggtctgaca acagcttgcc aaatttggtt agtccttgga tcagagcctg 2563

ttgtgggagg tagggaggaa atatgtaaag aaaaacagga agatacctgc actaatcatt 2623

cagacttcat tgagctctgc aaactttgcc tggttgctat tggctacctt gatttgaaat 2683

gctttgtgaa aaaaggcact tttaacatca tagccacaga aatcaagtgc cagtctatct 2743

ggaatccatg ttgtattgca gataatgttc tcatttattt ttg 2786

<210> 14  
<211> 738  
<212> PRT  
<213> Unknown

<400> 14  
Met Ala Pro Trp Leu Gln Leu Cys Ser Val Phe Phe Thr Val Asn Ala  
-15 -10 -5 -1

Cys Leu Asn Gly Ser Gln Leu Ala Xaa Ala Ala Gly Gly Ser Gly Arg  
 1 5 10 15  
 Ala Xaa Gly Ala Asp Thr Cys Ser Trp Xaa Gly Val Gly Pro Ala Ser  
 20 25 30  
 Arg Asn Ser Gly Leu Tyr Asn Ile Thr Phe Lys Tyr Asp Asn Cys Thr  
 35 40 45  
 Thr Tyr Leu Asn Pro Val Gly Lys His Val Ile Ala Asp Ala Gln Asn  
 50 55 60  
 Ile Thr Ile Ser Gln Tyr Ala Cys His Asp Gln Val Ala Val Thr Ile  
 65 70 75 80  
 Leu Trp Ser Pro Gly Ala Leu Gly Ile Glu Phe Leu Lys Gly Phe Arg  
 85 90 95  
 Val Ile Leu Glu Glu Leu Lys Ser Glu Gly Arg Gln Xaa Gln Gln Leu  
 100 105 110  
 Ile Leu Lys Asp Pro Lys Gln Xaa Asn Ser Ser Phe Lys Arg Thr Gly  
 115 120 125  
 Met Glu Ser Gln Pro Xaa Leu Asn Met Lys Phe Glu Thr Asp Tyr Phe  
 130 135 140  
 Val Arg Leu Ser Phe Ser Phe Ile Lys Asn Glu Ser Asn Tyr His Pro  
 145 150 155 160  
 Phe Phe Phe Arg Thr Arg Ala Cys Asp Leu Leu Leu Gln Pro Asp Asn  
 165 170 175  
 Leu Ala Cys Lys Pro Phe Trp Lys Pro Arg Asn Leu Asn Ile Ser Gln  
 180 185 190  
 His Gly Ser Asp Met Gln Val Ser Phe Asp His Ala Pro His Asn Phe  
 195 200 205  
 Gly Phe Arg Phe Phe Tyr Leu His Tyr Lys Leu Lys His Glu Gly Pro  
 210 215 220  
 Phe Lys Arg Lys Thr Cys Lys Gln Glu Gln Thr Thr Glu Met Thr Ser  
 225 230 235 240  
 Cys Leu Leu Gln Asn Val Ser Pro Gly Asp Tyr Ile Ile Glu Leu Val  
 245 250 255  
 Asp Asp Thr Asn Thr Thr Arg Lys Val Met His Tyr Ala Leu Lys Pro  
 260 265 270  
 Val His Ser Pro Trp Ala Gly Pro Ile Arg Ala Val Ala Ile Thr Val  
 275 280 285  
 Pro Leu Val Val Ile Ser Ala Phe Ala Thr Leu Phe Thr Val Met Cys  
 290 295 300  
 Arg Lys Lys Gln Gln Glu Asn Ile Tyr Ser His Leu Asp Glu Glu Ser  
 305 310 315 320



Ser Glu Ser Ser Thr Tyr Thr Ala Ala Leu Pro Arg Glu Arg Leu Arg  
 325 330 335  
 Pro Arg Pro Lys Val Phe Leu Cys Tyr Ser Ser Lys Asp Gly Gln Asn  
 340 345 350  
 His Met Asn Val Val Gln Cys Phe Ala Tyr Phe Leu Gln Asp Phe Cys  
 355 360 365  
 Gly Cys Glu Val Ala Leu Asp Leu Trp Glu Asp Phe Ser Leu Cys Arg  
 370 375 380  
 Glu Gly Gln Arg Glu Trp Val Ile Gln Lys Ile His Glu Ser Gln Phe  
 385 390 395 400  
 Ile Ile Val Val Cys Ser Lys Gly Met Lys Tyr Phe Val Asp Lys Lys  
 405 410 415  
 Asn Tyr Lys His Lys Gly Gly Gly Arg Gly Ser Gly Lys Gly Glu Leu  
 420 425 430  
 Phe Leu Val Ala Val Ser Ala Ile Ala Glu Lys Leu Arg Gln Ala Lys  
 435 440 445  
 Gln Ser Ser Ser Ala Ala Leu Ser Lys Phe Ile Ala Val Tyr Phe Asp  
 450 455 460  
 Tyr Ser Cys Glu Gly Asp Val Pro Gly Ile Leu Asp Leu Ser Thr Lys  
 465 470 475 480  
 Tyr Arg Leu Met Asp Asn Leu Pro Gln Leu Cys Ser His Leu His Ser  
 485 490 495  
 Arg Asp His Gly Leu Gln Glu Pro Gly Gln His Thr Arg Gln Gly Ser  
 500 505 510  
 Arg Arg Asn Tyr Phe Arg Ser Lys Ser Gly Arg Ser Leu Tyr Val Ala  
 515 520 525  
 Ile Cys Asn Met His Gln Phe Ile Asp Glu Glu Pro Asp Trp Phe Glu  
 530 535 540  
 Lys Gln Phe Val Pro Phe His Pro Pro Pro Leu Arg Tyr Arg Glu Pro  
 545 550 555 560  
 Val Leu Glu Lys Phe Asp Ser Gly Leu Val Leu Asn Asp Val Met Cys  
 565 570 575  
 Lys Pro Gly Pro Glu Ser Asp Phe Cys Leu Lys Val Glu Ala Ala Val  
 580 585 590  
 Leu Gly Ala Thr Gly Pro Ala Asp Ser Gln His Glu Ser Gln His Gly  
 595 600 605  
 Gly Leu Asp Gln Asp Gly Glu Ala Arg Pro Ala Leu Asp Gly Ser Ala  
 610 615 620  
 Ala Leu Gln Pro Leu Leu His Thr Val Lys Ala Gly Ser Pro Ser Asp  
 625 630 635 640

Met Pro Arg Asp Ser Gly Ile Tyr Asp Ser Ser Val Pro Ser Ser Glu  
645 650 655

Leu Ser Leu Pro Leu Met Glu Gly Leu Ser Thr Asp Gln Thr Glu Thr  
660 665 670

Ser Ser Leu Thr Glu Ser Val Ser Ser Ser Ser Gly Leu Gly Glu Glu  
675 680 685

Glu Pro Pro Ala Leu Pro Ser Lys Leu Leu Ser Ser Gly Ser Cys Lys  
690 695 700

Ala Asp Leu Gly Cys Arg Ser Tyr Thr Asp Glu Leu His Ala Val Ala  
705 710 715 720

Pro Leu

<210> 15  
<211> 2214  
<212> DNA  
<213> reverse translation

<220>  
<221> misc\_feature  
<222> (1)..(2214)  
<223> n may be a, c, g, or t

<400> 15  
atggcncnt ggytncaryt ntgywsngtn ttyttyacng tnaaygcntg yytnaayggn 60  
wsncarytng cngtngcngc nggnngnwsn ggnmgngcnn nngngncnga yacntgywsn 120  
tggnnnngng tnggncncgc nwsnmgnaay wsnggnytn tayaayathac nttyaartay 180  
gayaaytgaa cnacntayyt naayccngtn ggnaarcayg tnathgcnga ygcncaraay 240  
athacnathw sncartaygc ntgycaayg cargtngcng tnacnathyt ntggwsncn 300  
ggngcnytn gnatgcartt yytnaarggn ttmngntna thytngarga rytnaarwsn 360  
garggnmgnc arnnncarca rytathyt aargayccna arcarnnnaa ywsnwsntty 420  
aarmgnacng gnatggarws ncarccnnnn ytnaayatga arttygarac ngaytaytty 480  
gtmngnytnw snttywsntt yathaaraay garwsnaayt aycayccntt yttyttymgn 540  
acnmngcncnt gygayytnyt nytnarcncn gayaayytng cntgyaarcc nttytggaar 600  
ccnmgnaayy tnaayathws ncarcayggn wsngayatgc argtnwsntt ygaycaygcn 660  
ccncayaayt tyggnttymg nttytttyay ytncaytaya arytnaarca ygarggncn 720  
ttyaarmgna aracntgyaa rcargarcar acnacngara tgacnwsntg yytnytnear 780  
aaygtnwsnc cngngayta yathathgar ytngtngayg ayacnaayac nacnmgnaar 840  
gtnatgcayt aygcnytnaa rccngtncay wsncntggg cnggncnat hmgngcngtn 900

gcnathacng tnccnytngt ngtnathwsn gcnttygcna cnytnntyac ngtnatgtgy 960  
 mgnaaraarc arcargaraa yathtaywsn cayytngayg argarwsnws ngarwsnwsn 1020  
 acntayacng cngcnytncc nmngngarmgn ytnmgncnm gncnaargt nttyytnngy 1080  
 taywsnwsna argayggna raaycaytg aaygtngtnc artgyttygc ntayttytn 1140  
 cargayttyp gyggntgyga rgtngcnytn gayytnnggg argayttyws nytntgymgn 1200  
 garggncarm gngartgggt nathcaraar athcaygarw sncarttyat hathgtngtn 1260  
 tgywsnaarg gnatgaarta yttygtngay aaraaraayt ayaarcayaa rggngggngn 1320  
 mgnggnwsng gnaarggnga rytnttytn gtngcngtnw sngcnathgc ngaraarytn 1380  
 mgncargcna arcarsnws nwsngcngcn ytnwsnaart tyathgcngt ntayttygay 1440  
 taywsntgyg arggngaygt nccnggnath ytngayytnw snacnaarta ymgnytnatg 1500  
 gayaayytn cncarytnng ywsncayytn caywsnmngn aycayggnyt ncargarccn 1560  
 ggnarcaya cnmgncargg nwsnmgnmgn aaytaytym gnwsnaarws nggnmgnwsn 1620  
 ytntaygtng cnathtgyaa yatgcaycar ttyathgayg argarccnga ytggttygar 1680  
 aarcarttyg tnccnttyca yccnccnccn ytnmgntaym gngarccngt nytngaraar 1740  
 ttygaywsng gnytnngtnt naaygaygt atgtgyaarc cnggnccnga rwsngaytty 1800  
 tgyytnaarg tngargcngc ngtnytnngn gcnacnggnc cngcngayws ncarcaygar 1860  
 wsnarcayg gnggnytna ycargaygn gargcnmgnc cngcnytna yggnwngcn 1920  
 gcnytnarc cnytnytna yacngtnaar gcnggnwsnc cnwsngayat gccnmngay 1980  
 wsggnatht aygaywsnws ngtnccnwsn wsngarytnw snytnccnyt natggarggn 2040  
 ytnwsnacng aycaracnga racnwsnwsn ytnacngarw sngtnwsnws nwsnwsngn 2100  
 ytnngngarg argarccncc ngcnytnccn wsnaarytny tnwsnwsngg nwsntgyaar 2160  
 gcngayytn gntgymgnws ntayacngay garytncayg cngtngcncc nytn 2214

<210> 16

<211> 2012

<212> DNA

<213> Unknown

<220>

<223> Description of Unknown Organism: primate; surmised  
Homo sapiens

<220>

<221> CDS

<222> (1)..(1971)

<220>

&lt;221&gt; mat\_peptide

&lt;222&gt; (70)..(1971)

&lt;400&gt; 16

atg ggg agc tcc aga ctg gca gcc ctg ctc ctg cct ctc ctc ctc ata	48
Met Gly Ser Ser Arg Leu Ala Ala Leu Leu Leu Pro Leu Leu Leu Ile	
-20 -15 -10	
gtc atc gac ctc tct gac tct gct ggg att ggc ttt cgc cac ctg ccc	96
Val Ile Asp Leu Ser Asp Ser Ala Gly Ile Gly Phe Arg His Leu Pro	
-5 -1 1 5	
cac tgg aac acc cgc tgt cct ctg gcc tcc cac acg gaa gtt ctg cct	144
His Trp Asn Thr Arg Cys Pro Leu Ala Ser His Thr Glu Val Leu Pro	
10 15 20 25	
ata tcc ctt gcc gca cct ggt ggg ccc tct tct cca caa agc ctt ggt	192
Ile Ser Leu Ala Ala Pro Gly Gly Pro Ser Ser Pro Gln Ser Leu Gly	
30 35 40	
gtg tgc gag tct ggc act gtt ccc gct gtt tgt gcc agc atc tgc tgt	240
Val Cys Glu Ser Gly Thr Val Pro Ala Val Cys Ala Ser Ile Cys Cys	
45 50 55	
cag gtg gct cag gtc ttc aac ggg gcc tct tcc acc tcc tgg tgc aga	288
Gln Val Ala Gln Val Phe Asn Gly Ala Ser Ser Thr Ser Trp Cys Arg	
60 65 70	
aat cca aaa agt ctt cca cat tca agt tct ata gga gac aca aga tgc	336
Asn Pro Lys Ser Leu Pro His Ser Ser Ser Ile Gly Asp Thr Arg Cys	
75 80 85	
cag cac ctg ctc aga gga agc tgc tgc ctc gtc gtc acc tgt ctg aga	384
Gln His Leu Leu Arg Gly Ser Cys Cys Leu Val Val Thr Cys Leu Arg	
90 95 100 105	
aga gcc atc aca ttt cca tcc cct ccc cag aca tct ccc aca agg gac	432
Arg Ala Ile Thr Phe Pro Ser Pro Pro Gln Thr Ser Pro Thr Arg Asp	
110 115 120	
ttc gct cta aaa gga ccc aac ctt cgg atc cag aga cat ggg aaa gtc	480
Phe Ala Leu Lys Gly Pro Asn Leu Arg Ile Gln Arg His Gly Lys Val	
125 130 135	
ttc cca gat tgg act cac aaa ggc atg gag gtg ggc act ggg tac aac	528
Phe Pro Asp Trp Thr His Lys Gly Met Glu Val Gly Thr Gly Tyr Asn	
140 145 150	
agg aga tgg gtt cag ctg agt ggt gga ccc gag ttc tcc ttt gat ttg	576
Arg Arg Trp Val Gln Leu Ser Gly Gly Pro Glu Phe Ser Phe Asp Leu	
155 160 165	
ctg cct gag gcc cgg gct att cgg gtg acc ata tct tca ggc cct gag	624
Leu Pro Glu Ala Arg Ala Ile Arg Val Thr Ile Ser Ser Gly Pro Glu	
170 175 180 185	
gtc agc gtg cgt ctt tgt cac cag tgg gca ctg gag tgt gaa gag ctg	672
Val Ser Val Arg Leu Cys His Gln Trp Ala Leu Glu Cys Glu Glu Leu	
190 195 200	

agc agt ccc tat gat gtc cag aaa att gtg tct ggg ggc cac act gta	720
Ser Ser Pro Tyr Asp Val Gln Lys Ile Val Ser Gly Gly His Thr Val	
205 210 215	
gag ctg cct tat gaa ttc ctt ctg ccc tgt ctg tgc ata gag gca tcc	768
Glu Leu Pro Tyr Glu Phe Leu Leu Pro Cys Leu Cys Ile Glu Ala Ser	
220 225 230	
tac ctg caa gag gac act gtg agg cgc aaa aaa tgt ccc ttc cag agc	816
Tyr Leu Gln Glu Asp Thr Val Arg Arg Lys Lys Cys Pro Phe Gln Ser	
235 240 245	
tgg cca gaa gcc tat ggc tcg gac ttc tgg aag tca gtg cac ttc act	864
Trp Pro Glu Ala Tyr Ser Asp Phe Trp Lys Ser Val His Phe Thr	
250 255 260 265	
gac tac agc cag cac act cag atg gtc atg gcc ctg aca ctc cgc tgc	912
Asp Tyr Ser Gln His Thr Gln Met Val Met Ala Leu Thr Leu Arg Cys	
270 275 280	
cca ctg aag ctg gaa gct gcc ctc tgc cag agg cac gac tgg cat acc	960
Pro Leu Lys Leu Glu Ala Ala Leu Cys Gln Arg His Asp Trp His Thr	
285 290 295	
ctt tgc aaa gac ctc ccg aat gcc acg gct cga gag tca gat ggg tgg	1008
Leu Cys Lys Asp Leu Pro Asn Ala Thr Ala Arg Glu Ser Asp Gly Trp	
300 305 310	
tat gtt ttg gag aag gtg gac ctg cac ccc cag ctc tgc ttc aag gta	1056
Tyr Val Leu Glu Lys Val Asp Leu His Pro Gln Leu Cys Phe Lys Val	
315 320 325	
caa cca tgg ttc tct ttt gga aac agc agc cat gtt gaa tgc ccc cac	1104
Gln Pro Trp Phe Ser Phe Gly Asn Ser Ser His Val Glu Cys Pro His	
330 335 340 345	
cag act ggg tct ctc aca tcc tgg aat gta agc atg gat acc caa gcc	1152
Gln Thr Gly Ser Leu Thr Ser Trp Asn Val Ser Met Asp Thr Gln Ala	
350 355 360	
cag cag ctg att ctt cac ttc tcc tca aga atg cat gcc acc ttc agt	1200
Gln Gln Leu Ile Leu His Phe Ser Ser Arg Met His Ala Thr Phe Ser	
365 370 375	
gct gcc tgg agc ctc cca ggc ttg ggg cag gac act ttg gtg ccc ccc	1248
Ala Ala Trp Ser Leu Pro Gly Leu Gly Gln Asp Thr Leu Val Pro Pro	
380 385 390	
gtg tac act gtc agc cag gtg tgg cgg tca gat gtc cag ttt gcc tgg	1296
Val Tyr Thr Val Ser Gln Val Trp Arg Ser Asp Val Gln Phe Ala Trp	
395 400 405	
aag cac ctc ttg tgt cca gat gtc tct tac aga cac ctg ggg ctc ttg	1344
Lys His Leu Leu Cys Pro Asp Val Ser Tyr Arg His Leu Gly Leu Leu	
410 415 420 425	
atc ctg gca ctg ctg gcc ctc ctc acc cta ctg ggt gtt gtt ctg gcc	1392
Ile Leu Ala Leu Leu Ala Leu Leu Thr Leu Leu Gly Val Val Leu Ala	
430 435 440	

ctc acc tgc cgg cgc cca cag tca ggc ccg ggc cca gcg cgg cca gtg 1440  
 Leu Thr Cys Arg Arg Pro Gln Ser Gly Pro Gly Pro Ala Arg Pro Val  
 445 450 455

ctc ctc ctg cac gcg gcg gac tcg gag gcg cag cgg cgc ctg gtg gga 1488  
 Leu Leu Leu His Ala Ala Asp Ser Glu Ala Gln Arg Arg Leu Val Gly  
 460 465 470

gcg ctg gct gaa ctg cta cgg gca gcg ctg ggc ggc ggg cgc gac gtg 1536  
 Ala Leu Ala Glu Leu Leu Arg Ala Ala Leu Gly Gly Gly Arg Asp Val  
 475 480 485

atc gtg gac ctg tgg gag ggg agg cac gtg gcg cgc gtg ggc ccg ctg 1584  
 Ile Val Asp Leu Trp Glu Gly Arg His Val Ala Arg Val Gly Pro Leu  
 490 495 500 505

ccg tgg ctc tgg gcg gcg cgg acg cgc gta gcg cgg gag cag ggc act 1632  
 Pro Trp Leu Trp Ala Ala Arg Thr Arg Val Ala Arg Glu Gln Gly Thr  
 510 515 520

gtg ctg ctg ctg tgg agc ggc gcc gac ctt cgc ccg gtc agc ggc ccc 1680  
 Val Leu Leu Leu Trp Ser Gly Ala Asp Leu Arg Pro Val Ser Gly Pro  
 525 530 535

gac ccc cgc gcc gcg ccc ctg ctc gcc ctg ctc cac gct gcc ccg cgc 1728  
 Asp Pro Arg Ala Ala Pro Leu Leu Ala Leu Leu His Ala Ala Pro Arg  
 540 545 550

ccg ctg ctg ctg ctc gct tac ttc agt cgc ctc tgc gcc aag ggc gac 1776  
 Pro Leu Leu Leu Leu Ala Tyr Phe Ser Arg Leu Cys Ala Lys Gly Asp  
 555 560 565

atc ccc ccg ccg ctg cgc gcc ctg ccg cgc tac cgc ctg ctg cgc gac 1824  
 Ile Pro Pro Pro Leu Arg Ala Leu Pro Arg Tyr Arg Leu Leu Arg Asp  
 570 575 580 585

ctg ccg cgt ctg ctg cgg gcg ctg gac gcg cgg cct ttc gca gag gcc 1872  
 Leu Pro Arg Leu Leu Arg Ala Leu Asp Ala Arg Pro Phe Ala Glu Ala  
 590 595 600

acc agc tgg ggc cgc ctt ggg gcg cgg cag cgc agg cag agc cgc cta 1920  
 Thr Ser Trp Gly Arg Leu Gly Ala Arg Gln Arg Arg Gln Ser Arg Leu  
 605 610 615

gag ctg tgc agc cgg ctc gaa cga gag gcc gcc cga ctt gca gac cta 1968  
 Glu Leu Cys Ser Arg Leu Glu Arg Glu Ala Ala Arg Leu Ala Asp Leu  
 620 625 630

ggt tgagcagagc tccaccgcag tcccgggtgt ctgcggccgc t 2012  
 Gly

&lt;210&gt; 17

&lt;211&gt; 657

&lt;212&gt; PRT

&lt;213&gt; Unknown

&lt;400&gt; 17

Met Gly Ser Ser Arg Leu Ala Ala Leu Leu Leu Pro Leu Leu Leu Ile  
 -20 -15 -10

Val Ile Asp Leu Ser Asp Ser Ala Gly Ile Gly Phe Arg His Leu Pro  
                   -5                  -1  1                  5  
 His Trp Asn Thr Arg Cys Pro Leu Ala Ser His Thr Glu Val Leu Pro  
 10                                  15                                  20                                  25  
 Ile Ser Leu Ala Ala Pro Gly Gly Pro Ser Ser Pro Gln Ser Leu Gly  
                                   30                                  35                                  40  
 Val Cys Glu Ser Gly Thr Val Pro Ala Val Cys Ala Ser Ile Cys Cys  
                                   45                                  50                                  55  
 Gln Val Ala Gln Val Phe Asn Gly Ala Ser Ser Thr Ser Trp Cys Arg  
                                   60                                  65                                  70  
 Asn Pro Lys Ser Leu Pro His Ser Ser Ser Ile Gly Asp Thr Arg Cys  
                                   75                                  80                                  85  
 Gln His Leu Leu Arg Gly Ser Cys Cys Leu Val Val Thr Cys Leu Arg  
 90                                  95                                  100                                  105  
 Arg Ala Ile Thr Phe Pro Ser Pro Pro Gln Thr Ser Pro Thr Arg Asp  
                                   110                                  115                                  120  
 Phe Ala Leu Lys Gly Pro Asn Leu Arg Ile Gln Arg His Gly Lys Val  
                                   125                                  130                                  135  
 Phe Pro Asp Trp Thr His Lys Gly Met Glu Val Gly Thr Gly Tyr Asn  
                                   140                                  145                                  150  
 Arg Arg Trp Val Gln Leu Ser Gly Gly Pro Glu Phe Ser Phe Asp Leu  
                                   155                                  160                                  165  
 Leu Pro Glu Ala Arg Ala Ile Arg Val Thr Ile Ser Ser Gly Pro Glu  
 170                                  175                                  180                                  185  
 Val Ser Val Arg Leu Cys His Gln Trp Ala Leu Glu Cys Glu Glu Leu  
                                   190                                  195                                  200  
 Ser Ser Pro Tyr Asp Val Gln Lys Ile Val Ser Gly Gly His Thr Val  
                                   205                                  210                                  215  
 Glu Leu Pro Tyr Glu Phe Leu Leu Pro Cys Leu Cys Ile Glu Ala Ser  
                                   220                                  225                                  230  
 Tyr Leu Gln Glu Asp Thr Val Arg Arg Lys Lys Cys Pro Phe Gln Ser  
                                   235                                  240                                  245  
 Trp Pro Glu Ala Tyr Gly Ser Asp Phe Trp Lys Ser Val His Phe Thr  
 250                                  255                                  260                                  265  
 Asp Tyr Ser Gln His Thr Gln Met Val Met Ala Leu Thr Leu Arg Cys  
                                   270                                  275                                  280  
 Pro Leu Lys Leu Glu Ala Ala Leu Cys Gln Arg His Asp Trp His Thr  
                                   285                                  290                                  295  
 Leu Cys Lys Asp Leu Pro Asn Ala Thr Ala Arg Glu Ser Asp Gly Trp  
                                   300                                  305                                  310

Tyr Val Leu Glu Lys Val Asp Leu His Pro Gln Leu Cys Phe Lys Val  
 315 320 325  
 Gln Pro Trp Phe Ser Phe Gly Asn Ser Ser His Val Glu Cys Pro His  
 330 335 340 345  
 Gln Thr Gly Ser Leu Thr Ser Trp Asn Val Ser Met Asp Thr Gln Ala  
 350 355 360  
 Gln Gln Leu Ile Leu His Phe Ser Ser Arg Met His Ala Thr Phe Ser  
 365 370 375  
 Ala Ala Trp Ser Leu Pro Gly Leu Gly Gln Asp Thr Leu Val Pro Pro  
 380 385 390  
 Val Tyr Thr Val Ser Gln Val Trp Arg Ser Asp Val Gln Phe Ala Trp  
 395 400 405  
 Lys His Leu Leu Cys Pro Asp Val Ser Tyr Arg His Leu Gly Leu Leu  
 410 415 420 425  
 Ile Leu Ala Leu Leu Ala Leu Leu Thr Leu Leu Gly Val Val Leu Ala  
 430 435 440  
 Leu Thr Cys Arg Arg Pro Gln Ser Gly Pro Gly Pro Ala Arg Pro Val  
 445 450 455  
 Leu Leu Leu His Ala Ala Asp Ser Glu Ala Gln Arg Arg Leu Val Gly  
 460 465 470  
 Ala Leu Ala Glu Leu Leu Arg Ala Ala Leu Gly Gly Gly Arg Asp Val  
 475 480 485  
 Ile Val Asp Leu Trp Glu Gly Arg His Val Ala Arg Val Gly Pro Leu  
 490 495 500 505  
 Pro Trp Leu Trp Ala Ala Arg Thr Arg Val Ala Arg Glu Gln Gly Thr  
 510 515 520  
 Val Leu Leu Leu Trp Ser Gly Ala Asp Leu Arg Pro Val Ser Gly Pro  
 525 530 535  
 Asp Pro Arg Ala Ala Pro Leu Leu Ala Leu Leu His Ala Ala Pro Arg  
 540 545 550  
 Pro Leu Leu Leu Leu Ala Tyr Phe Ser Arg Leu Cys Ala Lys Gly Asp  
 555 560 565  
 Ile Pro Pro Pro Leu Arg Ala Leu Pro Arg Tyr Arg Leu Leu Arg Asp  
 570 575 580 585  
 Leu Pro Arg Leu Leu Arg Ala Leu Asp Ala Arg Pro Phe Ala Glu Ala  
 590 595 600  
 Thr Ser Trp Gly Arg Leu Gly Ala Arg Gln Arg Arg Gln Ser Arg Leu  
 605 610 615  
 Glu Leu Cys Ser Arg Leu Glu Arg Glu Ala Ala Arg Leu Ala Asp Leu  
 620 625 630



Gly

<210> 18  
<211> 1971  
<212> DNA  
<213> reverse translation  
  
<220>  
<221> misc\_feature  
<222> (1)..(1971)  
<223> n may be a, c, g, or t

<400> 18  
atgggnwsnw snmgnytn gc ngcnytnytn ytnccnytny tnytnathgt nathgayytn 60  
wsngaywsng cnggnathgg nttymgncay ytnccncayt ggaayacnmg ntgyccnytn 120  
gcwnsncaaya cngargtnyt nccnathwsn ytngcngcnc cnggnggncc nwsnwsnccn 180  
carwsnytn gngtntgyga rwsnggnacn gtncncngcng tntgygcnws nathtgytgy 240  
cargtngcnc argtnttyaa yggngcnwsn wsnacnwsnt ggtgymgnaa yccnaarwsn 300  
ytncncncayw snwsnwsnat hggngayacn mgntgyarc ayytnytnmg nggnwsntgy 360  
tgyytngtng tnaentgyt nmgnmgngcn athacnttyc cnwsnccncc ncaracnwsn 420  
ccnacnmng aytttygcny naarggnccn aaytnmgna thcarmgna yggnaargtn 480  
ttyccngayt ggacncayaa rggnatggar gtnggnacng gntayaaymg nmngtgggt 540  
carytnwsng gnggnccnga rttysntty gayytnytn cngargcnmg ngcnathmgn 600  
gtnacnathw snwsnggncc ngargtnwsn gtnmgnytn gycaycartg ggcnytn gar 660  
tgygargary tnwsnwsncc ntaygaytn caraarathg tnwsngngng ncayacngtn 720  
garytnccnt aygarttyt nytnccntgy ytntgyathg argcnwsnta yytnccargar 780  
gayacngtnm gnmgnaaraa rtgyccntty carwsntggc cngargcnta yggnwsngay 840  
ttypggaarw sngtncaytt yacngaytay wsnarcaya cncaratggt natggcnytn 900  
acnytnmgnt gyccnytnaa rytngargcn gcnytnngyc armgncayga ytggcayacn 960  
ytntgyaarg ayytnccnaa ygcncngcn mgngarwsng ayggntggta ygtnytn gar 1020  
aargtngayy tncayccna rytntgyt aargtnarc cntggtyws nttyggnaay 1080  
wsnwsncayg tngartgycc ncaycaracn ggnwsnytna cnwsntggaa ygtwnsnatg 1140  
gayacncarg cncarcaryt nathytnca tytwsnwsnm gnatgcaygc nacnttywn 1200  
gcngcntggw snytnccngg nytnngncar gayacnytn tncncncngt ntayacngtn 1260  
wsncargtn ggmgnwsnga ygtncartty gcntggaarc ayytnytnng yccngaygn 1320

wsntaymgnc ayytnggnyt nytnathytn gcnynytng cnytnytnac nytnytnngn 1380  
 gtngtnytng cnytnacntg ymgngmgnccn carwsnggnc cnggncngc nmgnccngtn 1440  
 ytynytnyc aygcngcnga ywsngargcn carmgmngny tngtnggngc nytnngcgar 1500  
 ytynytnmgng cngcnytnng nggnggmgn gaygtathg tngayytnng ggarggmgn 1560  
 caygtngcnm gngtnggncc nytnccntgg ytntgggngc cmgnacnmg ngtngcnmgn 1620  
 garcarggna cngtnytnyt nytnnggwsn ggngcngayy tnmgnccngt nwsnggnccn 1680  
 gayccnmng cngcncnny nytnngcnytn ytncaygng cncnmgnc nytnytnytn 1740  
 ytngcntayt tywsnmgnyt ntgygcnaar ggngayathc cncncnny nmngcnytn 1800  
 ccnmngntaym gnytnytnmg ngayytnccn mgnynytnm gngcnytna ygcnmgnccn 1860  
 ttygcngarg cnacwsntg gggnmgnytn ggngcnmgnc armgmngnca rwsnmgnytn 1920  
 garytnngyw snmgnytna rmngargcn gcnmgnytn cngayytnng n 1971

<210> 19

<211> 808

<212> DNA

<213> Unknown

<220>

<223> Description of Unknown Organism:rodent; surmised  
 Mus musculus

<220>

<221> CDS

<222> (78)..(806)

<220>

<221> mat\_peptide

<222> (147)..(806)

<400> 19

cagctccggg ccagccctg ctgccctctt gcagacagga aagacatggt ctctgcgccc 60  
 tgatcctaca gaagctc atg ggg agc ccc aga ctg gca gcc ttg ctc ctg 110  
 Met Gly Ser Pro Arg Leu Ala Ala Leu Leu Leu  
 -20 -15  
 tct ctc ccg cta ctg ctc atc ggc ctc gct gtg tct gct cgg gtt gcc 158  
 Ser Leu Pro Leu Leu Leu Ile Gly Leu Ala Val Ser Ala Arg Val Ala  
 -10 -5 -1 1  
 tgc ccc tgc ctg cgg agt tgg acc agc cac tgt ctc ctg gcc tac cgt 206  
 Cys Pro Cys Leu Arg Ser Trp Thr Ser His Cys Leu Leu Ala Tyr Arg  
 5 10 15 20  
 gtg gat aaa cgt ttt gct ggc ctt cag tgg ggc tgg ttc cct ctc ttg 254  
 Val Asp Lys Arg Phe Ala Gly Leu Gln Trp Gly Trp Phe Pro Leu Leu  
 25 30 35  
 gtg agg aaa tct aaa agt cct cct aaa ttt gaa gac tat tgg agg cac 302

Val	Arg	Lys	Ser	Lys	Ser	Pro	Pro	Lys	Phe	Glu	Asp	Tyr	Trp	Arg	His		
			40					45					50				
agg	aca	cca	gca	tcc	ttc	cag	agg	aag	ctg	cta	ggc	agc	cct	tcc	ctg	350	
Arg	Thr	Pro	Ala	Ser	Phe	Gln	Arg	Lys	Leu	Leu	Gly	Ser	Pro	Ser	Leu		
		55					60					65					
tct	gag	gaa	agc	cat	cga	att	tcc	atc	ccc	tcc	tca	gcc	atc	tcc	cac	398	
Ser	Glu	Glu	Ser	His	Arg	Ile	Ser	Ile	Pro	Ser	Ser	Ala	Ile	Ser	His		
	70					75					80						
aga	ggc	caa	cgc	acc	aaa	agg	gcc	cag	cct	tca	gct	gca	gaa	gga	aga	446	
Arg	Gly	Gln	Arg	Thr	Lys	Arg	Ala	Gln	Pro	Ser	Ala	Ala	Glu	Gly	Arg		
	85				90					95					100		
gaa	cat	ctc	cct	gaa	gca	ggg	tca	caa	aag	tgt	gga	gga	cct	gaa	ttc	494	
Glu	His	Leu	Pro	Glu	Ala	Gly	Ser	Gln	Lys	Cys	Gly	Gly	Pro	Glu	Phe		
				105				110						115			
tcc	ttt	gat	ttg	ctg	ccc	gag	gtg	cag	gct	gtt	cgg	gtg	act	att	cct	542	
Ser	Phe	Asp	Leu	Leu	Pro	Glu	Val	Gln	Ala	Val	Arg	Val	Thr	Ile	Pro		
			120					125					130				
gca	ggc	ccc	aag	gca	cgt	gtg	cgc	ctt	tgt	tat	cag	tgg	gca	ctg	gaa	590	
Ala	Gly	Pro	Lys	Ala	Arg	Val	Arg	Leu	Cys	Tyr	Gln	Trp	Ala	Leu	Glu		
		135				140						145					
tgt	gaa	gac	ttg	agt	agc	cct	ttt	gat	acc	cag	aaa	att	gtg	tct	gga	638	
Cys	Glu	Asp	Leu	Ser	Ser	Pro	Phe	Asp	Thr	Gln	Lys	Ile	Val	Ser	Gly		
	150					155					160						
ggg	cac	act	gta	gac	ctg	cct	tat	gaa	ttc	ctt	ctg	ccc	tgc	atg	tgc	686	
Gly	His	Thr	Val	Asp	Leu	Pro	Tyr	Glu	Phe	Leu	Leu	Pro	Cys	Met	Cys		
	165				170					175					180		
ata	gag	gcc	tcc	tac	ctg	caa	gag	gac	act	gtg	agg	cgc	aaa	agt	gtc	734	
Ile	Glu	Ala	Ser	Tyr	Leu	Gln	Glu	Asp	Thr	Val	Arg	Arg	Lys	Ser	Val		
				185					190					195			
cct	tcc	aga	gct	ggc	ctg	aag	ctt	atg	gct	cag	act	tct	ggc	agt	caa	782	
Pro	Ser	Arg	Ala	Gly	Leu	Lys	Leu	Met	Ala	Gln	Thr	Ser	Gly	Ser	Gln		
			200					205					210				
tac	gct	tca	ctg	act	aca	gcc	agc	ac								808	
Tyr	Ala	Ser	Leu	Thr	Thr	Ala	Ser										
		215					220										
<210> 20																	
<211> 243																	
<212> PRT																	
<213> Unknown																	
<400> 20																	
Met	Gly	Ser	Pro	Arg	Leu	Ala	Ala	Leu	Leu	Leu	Ser	Leu	Pro	Leu	Leu		
			-20					-15					-10				
Leu	Ile	Gly	Leu	Ala	Val	Ser	Ala	Arg	Val	Ala	Cys	Pro	Cys	Leu	Arg		
		-5				-1	1				5						

```

Ser Trp Thr Ser His Cys Leu Leu Ala Tyr Arg Val Asp Lys Arg Phe
 10                      15                      20                      25

Ala Gly Leu Gln Trp Gly Trp Phe Pro Leu Leu Val Arg Lys Ser Lys
                      30                      35                      40

Ser Pro Pro Lys Phe Glu Asp Tyr Trp Arg His Arg Thr Pro Ala Ser
                      45                      50                      55

Phe Gln Arg Lys Leu Leu Gly Ser Pro Ser Leu Ser Glu Glu Ser His
                      60                      65                      70

Arg Ile Ser Ile Pro Ser Ser Ala Ile Ser His Arg Gly Gln Arg Thr
 75                      80                      85

Lys Arg Ala Gln Pro Ser Ala Ala Glu Gly Arg Glu His Leu Pro Glu
 90                      95                      100                      105

Ala Gly Ser Gln Lys Cys Gly Gly Pro Glu Phe Ser Phe Asp Leu Leu
                      110                      115                      120

Pro Glu Val Gln Ala Val Arg Val Thr Ile Pro Ala Gly Pro Lys Ala
                      125                      130                      135

Arg Val Arg Leu Cys Tyr Gln Trp Ala Leu Glu Cys Glu Asp Leu Ser
                      140                      145                      150

Ser Pro Phe Asp Thr Gln Lys Ile Val Ser Gly Gly His Thr Val Asp
                      155                      160                      165

Leu Pro Tyr Glu Phe Leu Leu Pro Cys Met Cys Ile Glu Ala Ser Tyr
                      170                      175                      180                      185

Leu Gln Glu Asp Thr Val Arg Arg Lys Ser Val Pro Ser Arg Ala Gly
                      190                      195                      200

Leu Lys Leu Met Ala Gln Thr Ser Gly Ser Gln Tyr Ala Ser Leu Thr
                      205                      210                      215

Thr Ala Ser
                      220

```

&lt;210&gt; 21

&lt;211&gt; 729

&lt;212&gt; DNA

&lt;213&gt; reverse translation

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1)..(729)

&lt;223&gt; n may be a, c, g, or t

&lt;400&gt; 21

atgggnwsnc cnmgnytngc ngcnytnytn ytawsnytn cnytnytnyt nathggnytn 60

gcngtnwsng cnmgngtngc ntgyccntgy ytnmgnwsnt ggacnwsnca ytggytnytn 120

gcntaymgng tngayaarmg nttygcnggn ytncartggg gntgggtycc nytnytnngtn 180

mgnaarwsna arwsnccncc naarttygar gaytaytggm gncaymgnac nccngcnwsn 240  
 ttycarmgna arytnytngg nwsnccnwsn ytnwsngarg arwsncaymg nathwsnath 300  
 ccnwsnwsng cnathwsnca ymgnggncar mgnacnaarm gngcncarcc nwsngcngcn 360  
 garggnmgng arcayytnc ngargcnggn wncaraart gyggnggncc ngarttywsn 420  
 ttygayytny tncngargt ncargcngtn mgngtnacna thcngcngg nccnaargcn 480  
 mgngtnmgny tntgytayca rtgggcnyn gartgygarg ayytnwsnws nccnttygay 540  
 acncaraara thgtnwsngg nggncayacn gtngayytnc cntaygartt yytnytncn 600  
 tgyatgtgya thgargcnws ntayytncar gargayacng tnmgnmgnaa rwsngtnccn 660  
 wsnmgngcng gnytnaaryt natggcncar acnwsnggnw sncartaygc nwsnytnacn 720  
 acngcnwsn 729

<210> 22

<211> 2377

<212> DNA

<213> Unknown

<220>

<223> Description of Unknown Organism: primate; surmised  
Homo sapiens

<220>

<221> CDS

<222> (180)..(1874)

<400> 22

ttttgagcag aggccttccta ggctccgtag aaatttgcat acagcttcca cttcctgctt 60  
 cagagcctgt tcttctactt acctgggccc ggagaagggtg gagggagacg agaagccgcc 120  
 gagagccgac taccctccgg gccagtcgtg tctgtccgtg gtggatctaa gaaactaga 179  
 atg aac cga agc att cct gtg gag gtt gat gaa tca gaa cca tac cca 227  
 Met Asn Arg Ser Ile Pro Val Glu Val Asp Glu Ser Glu Pro Tyr Pro  
 1 5 10 15  
 agt cag ttg ctg aaa cca atc cca gaa tat tcc ccg gaa gag gaa tca 275  
 Ser Gln Leu Leu Lys Pro Ile Pro Glu Tyr Ser Pro Glu Glu Glu Ser  
 20 25 30  
 gaa cca cct gct cca aat ata agg aac atg gca ccc aac agc ttg tct 323  
 Glu Pro Pro Ala Pro Asn Ile Arg Asn Met Ala Pro Asn Ser Leu Ser  
 35 40 45  
 gca ccc aca atg ctt cac aat tcc tcc gga gac ttt tct caa gct cac 371  
 Ala Pro Thr Met Leu His Asn Ser Ser Gly Asp Phe Ser Gln Ala His  
 50 55 60  
 tca acc ctg aaa ctt gca aat cac cag cgg cct gta tcc cgg cag gtc 419  
 Ser Thr Leu Lys Leu Ala Asn His Gln Arg Pro Val Ser Arg Gln Val  
 65 70 75 80

acc tgc ctg cgc act caa gtt ctg gag gac agt gaa gac agt ttc tgc	467
Thr Cys Leu Arg Thr Gln Val Leu Glu Asp Ser Glu Asp Ser Phe Cys	
85 90 95	
agg aga cac cca ggc ctg ggc aaa gct ttc cct tct ggg tgc tct gca	515
Arg Arg His Pro Gly Leu Gly Lys Ala Phe Pro Ser Gly Cys Ser Ala	
100 105 110	
gtc agc gag cct gcg tct gag tct gtg gtt gga gcc ctc cct gca gag	563
Val Ser Glu Pro Ala Ser Glu Ser Val Val Gly Ala Leu Pro Ala Glu	
115 120 125	
cat cag ttt tca ttt atg gaa aaa cgt aat caa tgg ctg gta tct cag	611
His Gln Phe Ser Phe Met Glu Lys Arg Asn Gln Trp Leu Val Ser Gln	
130 135 140	
ctt tca gcg gct tct cct gac act ggc cat gac tca gac aaa tca gac	659
Leu Ser Ala Ala Ser Pro Asp Thr Gly His Asp Ser Asp Lys Ser Asp	
145 150 155 160	
caa agt tta cct aat gcc tca gca gac tcc ttg ggc ggt agc cag gag	707
Gln Ser Leu Pro Asn Ala Ser Ala Asp Ser Leu Gly Gly Ser Gln Glu	
165 170 175	
atg gtg caa cgg ccc cag cct cac agg aac cga gca ggc ctg gat ctg	755
Met Val Gln Arg Pro Gln Pro His Arg Asn Arg Ala Gly Leu Asp Leu	
180 185 190	
cca acc ata gac acg gga tat gat tcc cag ccc cag gat gtc ctg ggc	803
Pro Thr Ile Asp Thr Gly Tyr Asp Ser Gln Pro Gln Asp Val Leu Gly	
195 200 205	
atc agg cag ctg gaa agg ccc ctg ccc ctc acc tcc gtg tgt tac ccc	851
Ile Arg Gln Leu Glu Arg Pro Leu Pro Leu Thr Ser Val Cys Tyr Pro	
210 215 220	
cag gac ctc ccc aga cct ctc agg tcc agg gag ttc cct cag ttt gaa	899
Gln Asp Leu Pro Arg Pro Leu Arg Ser Arg Glu Phe Pro Gln Phe Glu	
225 230 235 240	
cct cag agg tat cca gca tgt gca cag atg ctg cct ccc aat ctt tcc	947
Pro Gln Arg Tyr Pro Ala Cys Ala Gln Met Leu Pro Pro Asn Leu Ser	
245 250 255	
cca cat gct cca tgg aac tat cat tac cat tgt cct gga agt ccc gat	995
Pro His Ala Pro Trp Asn Tyr His Tyr His Cys Pro Gly Ser Pro Asp	
260 265 270	
cac cag gtg cca tat ggc cat gac tac cct cga gca gcc tac cag caa	1043
His Gln Val Pro Tyr Gly His Asp Tyr Pro Arg Ala Ala Tyr Gln Gln	
275 280 285	
gtg atc cag ccg gct ctg cct ggg cag ccc ctg cct gga gcc agt gtg	1091
Val Ile Gln Pro Ala Leu Pro Gly Gln Pro Leu Pro Gly Ala Ser Val	
290 295 300	
aga ggc ctg cac cct gtg cag aag gtt atc ctg aat tat ccc agc ccc	1139
Arg Gly Leu His Pro Val Gln Lys Val Ile Leu Asn Tyr Pro Ser Pro	
305 310 315 320	

tgg gac caa gaa gag agg ccc gca cag aga gac tgc tcc ttt ccg ggg	1187
Trp Asp Gln Glu Glu Arg Pro Ala Gln Arg Asp Cys Ser Phe Pro Gly	
325 330 335	
ctt cca agg cac cag gac cag cca cat cac cag cca cct aat aga gct	1235
Leu Pro Arg His Gln Asp Gln Pro His His Gln Pro Pro Asn Arg Ala	
340 345 350	
ggg gct cct ggg gag tcc ttg gag tgc cct gca gag ctg aga cca cag	1283
Gly Ala Pro Gly Glu Ser Leu Glu Cys Pro Ala Glu Leu Arg Pro Gln	
355 360 365	
gtt ccc cag cct ccg tcc cca gct gct gtg cct aga ccc cct agc aac	1331
Val Pro Gln Pro Pro Ser Pro Ala Ala Val Pro Arg Pro Pro Ser Asn	
370 375 380	
cct cca gcc aga gga act cta aaa aca agc aat ttg cca gaa gaa ttg	1379
Pro Pro Ala Arg Gly Thr Leu Lys Thr Ser Asn Leu Pro Glu Glu Leu	
385 390 395 400	
cgg aaa gtc ttt atc act tat tcg atg gac aca gct atg gag gtg gtg	1427
Arg Lys Val Phe Ile Thr Tyr Ser Met Asp Thr Ala Met Glu Val Val	
405 410 415	
aaa ttc gtg aac ttt ttg ttg gta aat ggc ttc caa act gca att gac	1475
Lys Phe Val Asn Phe Leu Leu Val Asn Gly Phe Gln Thr Ala Ile Asp	
420 425 430	
ata ttt gag gat aga atc cga ggc att gat atc att aaa tgg atg gag	1523
Ile Phe Glu Asp Arg Ile Arg Gly Ile Asp Ile Ile Lys Trp Met Glu	
435 440 445	
cgc tac ctt agg gat aag acc gtg atg ata atc gta gca atc agc ccc	1571
Arg Tyr Leu Arg Asp Lys Thr Val Met Ile Ile Val Ala Ile Ser Pro	
450 455 460	
aaa tac aaa cag gac gtg gaa ggc gct gag tcg cag ctg gac gag gat	1619
Lys Tyr Lys Gln Asp Val Glu Gly Ala Glu Ser Gln Leu Asp Glu Asp	
465 470 475 480	
gag cat ggc tta cat act aag tac att cat cga atg atg cag att gag	1667
Glu His Gly Leu His Thr Lys Tyr Ile His Arg Met Met Gln Ile Glu	
485 490 495	
ttc ata aaa caa gga agc atg aat ttc aga ttc atc cct gtg ctc ttc	1715
Phe Ile Lys Gln Gly Ser Met Asn Phe Arg Phe Ile Pro Val Leu Phe	
500 505 510	
cca aat gct aag aag gag cat gtg ccc acc tgg ctt cag aac act cat	1763
Pro Asn Ala Lys Lys Glu His Val Pro Thr Trp Leu Gln Asn Thr His	
515 520 525	
gtc tac agc tgg ccc aag aat aaa aaa aac atc ctg ctg cgg ctg ctg	1811
Val Tyr Ser Trp Pro Lys Asn Lys Lys Asn Ile Leu Leu Arg Leu Leu	
530 535 540	
aga gag gaa gag tat gtg gct cct cca cgg ggg cct ctg ccc acc ctt	1859
Arg Glu Glu Glu Tyr Val Ala Pro Pro Arg Gly Pro Leu Pro Thr Leu	
545 550 555 560	

cag gtg gtt ccc ttg tgacaccgtt catccccaga tcaactgaggc caggccatgt 1914  
 Gln Val Val Pro Leu  
 565

ttggggcctt gttctgacag cattctggct gaggctggtc ggtagcactc ctggctgggt 1974  
 tttttctggt cctccccgag aggcctcttg gccccagga aacctgttgt gcagagctct 2034  
 tccccggaga cctccacaca ccctggcttt gaagtggagt ctgtgactgc tctgcattct 2094  
 ctgcttttaa aaaaaccatt gcaggtgcc a gtgtcccata tgttcctcct gacagtttga 2154  
 tgtgtccatt ctgggcctct cagtgccttag caagtagata atgtaaggga tgtggcagca 2214  
 aatggaaatg actacaaaca ctctcctatc aatcacttca ggctactttt atgagttagc 2274  
 cagatgcttg tgtatcctca gaccaaactg attcatgtac aaataataaa atgtttactc 2334  
 ttttgtaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa 2377

<210> 23  
 <211> 565  
 <212> PRT  
 <213> Unknown

<400> 23  
 Met Asn Arg Ser Ile Pro Val Glu Val Asp Glu Ser Glu Pro Tyr Pro  
 1 5 10 15  
 Ser Gln Leu Leu Lys Pro Ile Pro Glu Tyr Ser Pro Glu Glu Glu Ser  
 20 25 30  
 Glu Pro Pro Ala Pro Asn Ile Arg Asn Met Ala Pro Asn Ser Leu Ser  
 35 40 45  
 Ala Pro Thr Met Leu His Asn Ser Ser Gly Asp Phe Ser Gln Ala His  
 50 55 60  
 Ser Thr Leu Lys Leu Ala Asn His Gln Arg Pro Val Ser Arg Gln Val  
 65 70 75 80  
 Thr Cys Leu Arg Thr Gln Val Leu Glu Asp Ser Glu Asp Ser Phe Cys  
 85 90 95  
 Arg Arg His Pro Gly Leu Gly Lys Ala Phe Pro Ser Gly Cys Ser Ala  
 100 105 110  
 Val Ser Glu Pro Ala Ser Glu Ser Val Val Gly Ala Leu Pro Ala Glu  
 115 120 125  
 His Gln Phe Ser Phe Met Glu Lys Arg Asn Gln Trp Leu Val Ser Gln  
 130 135 140  
 Leu Ser Ala Ala Ser Pro Asp Thr Gly His Asp Ser Asp Lys Ser Asp  
 145 150 155 160  
 Gln Ser Leu Pro Asn Ala Ser Ala Asp Ser Leu Gly Gly Ser Gln Glu  
 165 170 175



Met Val Gln Arg Pro Gln Pro His Arg Asn Arg Ala Gly Leu Asp Leu  
 180 185 190  
 Pro Thr Ile Asp Thr Gly Tyr Asp Ser Gln Pro Gln Asp Val Leu Gly  
 195 200 205  
 Ile Arg Gln Leu Glu Arg Pro Leu Pro Leu Thr Ser Val Cys Tyr Pro  
 210 215 220  
 Gln Asp Leu Pro Arg Pro Leu Arg Ser Arg Glu Phe Pro Gln Phe Glu  
 225 230 235 240  
 Pro Gln Arg Tyr Pro Ala Cys Ala Gln Met Leu Pro Pro Asn Leu Ser  
 245 250 255  
 Pro His Ala Pro Trp Asn Tyr His Tyr His Cys Pro Gly Ser Pro Asp  
 260 265 270  
 His Gln Val Pro Tyr Gly His Asp Tyr Pro Arg Ala Ala Tyr Gln Gln  
 275 280 285  
 Val Ile Gln Pro Ala Leu Pro Gly Gln Pro Leu Pro Gly Ala Ser Val  
 290 295 300  
 Arg Gly Leu His Pro Val Gln Lys Val Ile Leu Asn Tyr Pro Ser Pro  
 305 310 315 320  
 Trp Asp Gln Glu Glu Arg Pro Ala Gln Arg Asp Cys Ser Phe Pro Gly  
 325 330 335  
 Leu Pro Arg His Gln Asp Gln Pro His His Gln Pro Pro Asn Arg Ala  
 340 345 350  
 Gly Ala Pro Gly Glu Ser Leu Glu Cys Pro Ala Glu Leu Arg Pro Gln  
 355 360 365  
 Val Pro Gln Pro Pro Ser Pro Ala Ala Val Pro Arg Pro Pro Ser Asn  
 370 375 380  
 Pro Pro Ala Arg Gly Thr Leu Lys Thr Ser Asn Leu Pro Glu Glu Leu  
 385 390 395 400  
 Arg Lys Val Phe Ile Thr Tyr Ser Met Asp Thr Ala Met Glu Val Val  
 405 410 415  
 Lys Phe Val Asn Phe Leu Leu Val Asn Gly Phe Gln Thr Ala Ile Asp  
 420 425 430  
 Ile Phe Glu Asp Arg Ile Arg Gly Ile Asp Ile Ile Lys Trp Met Glu  
 435 440 445  
 Arg Tyr Leu Arg Asp Lys Thr Val Met Ile Ile Val Ala Ile Ser Pro  
 450 455 460  
 Lys Tyr Lys Gln Asp Val Glu Gly Ala Glu Ser Gln Leu Asp Glu Asp  
 465 470 475 480  
 Glu His Gly Leu His Thr Lys Tyr Ile His Arg Met Met Gln Ile Glu  
 485 490 495

Phe Ile Lys Gln Gly Ser Met Asn Phe Arg Phe Ile Pro Val Leu Phe  
                   500                                  505                                  510  
 Pro Asn Ala Lys Lys Glu His Val Pro Thr Trp Leu Gln Asn Thr His  
                   515                                  520                                  525  
 Val Tyr Ser Trp Pro Lys Asn Lys Lys Asn Ile Leu Leu Arg Leu Leu  
                   530                                  535                                  540  
 Arg Glu Glu Glu Tyr Val Ala Pro Pro Arg Gly Pro Leu Pro Thr Leu  
                   545                                  550                                  555                                  560  
 Gln Val Val Pro Leu  
                                   565

<210> 24  
 <211> 1695  
 <212> DNA  
 <213> reverse translation

<220>  
 <221> misc\_feature  
 <222> (1)..(1695)  
 <223> n may be a, c, g, or t

<400> 24  
 atgaaymgmw snathccngt ngargtngay garwsngarc cntayccnws ncarytnytn 60  
 aarccnathc cngartayws nccngargar garwsngarc cncngcncnc naayathmgn 120  
 aayatggcnc cnaaywsnyt nwsngcncnc acnatgytnc ayaaywsnws nggngaytty 180  
 wsnarcgnc aywsnacnyt naarytngcn aaycaycarm gncngtnws nmgnrcargtn 240  
 acntgyytnm gnacncargt nytngargay wsnrgargayw snttytgymg nmgnrcayccn 300  
 ggnytnngna argcnttycc nwsnggntgy wsnrgngtnw sngarccngc nwsngarwsn 360  
 gtngtnngng cnytnccngc ngarcaycar ttywsnttya tggaraarmg naaycartgg 420  
 ytngtnwsnc arytnwsngc ngcnwsnccn gayacnggnc aygaywsnga yaarwsngay 480  
 carwsnytn cnaaygcnws ngcngaywsn ytngngngnw sncargarat ggtncarmgn 540  
 ccncarccnc aymgnaaymg ngcnggnytn gayytnccna cnathgayac nggntaygay 600  
 wsnarcrcnc argaygtnyt nggnathmgn carytngarm gncnytncc nytnacnwsn 660  
 gtntgytayc ncargayyt nccnmgnccn ytnmgnwsnm gngarttycc ncarttygar 720  
 ccncarmgnt ayccngcntg ygncaratg ytnccncna ayytnwsncc ncaygcncn 780  
 tggaaaytayc ayaycaytg yccngnwsn ccngaycayc argtnccnta yggncaygay 840  
 tayccnmgng cngcntayca rcargtnath carccngcny tncngngnca rccnytnccn 900  
 ggngcnwsng tnmnggnytn ncayccngtn caraargtna thytnaayta yccnwsnccn 960

tgggaycarg argarmgncc ngcncarmgn gaytgywsnt tyccnggnynt nccnmgnacay 1020  
 cargaycarc cncaycayca rccnccnaay mgngcnggng cncnggnga rwsnytngar 1080  
 tgyccngcng arytnmgnc ncartnccn carccnccnw snccngcngc ngtnccnmgn 1140  
 ccnccnwsna ayccnccngc nmngngnacn ytnaaracnw snaaytncc ngargarytn 1200  
 mgnaargtnt tyathacnta ywsnatggay acngcnatgg argtngtnaa rttygtnaay 1260  
 ttyytnytny tnaayggntt ycaracngcn athgayatht tygargaymg nathmgnggn 1320  
 athgayatha thaartggat ggarmgntay ytnmgngaya aracngtnat gathathgtn 1380  
 gcnathwsnc cnaartayaa rcargaygtn garggngcng arwsncaryt ngaygargay 1440  
 garcayggny tncayacnaa rtayathcay mgnatgatgc arathgartt yathaarcar 1500  
 ggnwsnatga ayttymgntt yathccngtn ytnttyccna aygcnaaraa rgarcaygtn 1560  
 ccnacntggy tncaraayac ncaygtntay wsntggccna araayaaraa raayathytn 1620  
 ytnmgnytny tnmngngarga rgartaygtn gcncnccnm gnggncnynt nccnacnytn 1680  
 cargtngtnc cnytn 1695

<210> 25

<211> 1323

<212> DNA

<213> Unknown

<220>

<223> Description of Unknown Organism:rodent; surmised  
Mus musculus

<220>

<221> CDS

<222> (1)..(1026)

<400> 25

cag gac ctc cct ggg cct ctg agg tcc agg gaa ttg cca cct cag ttt	48
Gln Asp Leu Pro Gly Pro Leu Arg Ser Arg Glu Leu Pro Pro Gln Phe	
1 5 10 15	
gaa ctt gag agg tat cca atg aac gcc cag ctg ctg ccg ccc cat cct	96
Glu Leu Glu Arg Tyr Pro Met Asn Ala Gln Leu Leu Pro Pro His Pro	
20 25 30	
tcc cca cag gcc cca tgg aac tgt cag tac tac tgc ccc gga ggg ccc	144
Ser Pro Gln Ala Pro Trp Asn Cys Gln Tyr Tyr Cys Pro Gly Gly Pro	
35 40 45	
tac cac cac cag gtg cca cac ggc cat ggc tac cct cca gca gca gcc	192
Tyr His His Gln Val Pro His Gly His Gly Tyr Pro Pro Ala Ala Ala	
50 55 60	
tac cag caa gta ctc cag cct gct ctg cct ggg cag gtc ctt cct ggg	240
Tyr Gln Gln Val Leu Gln Pro Ala Leu Pro Gly Gln Val Leu Pro Gly	
65 70 75 80	

gca agg gca aga ggc cca cgc cct gtg cag aag gtc atc ctg aat gac	288
Ala Arg Ala Arg Gly Pro Arg Pro Val Gln Lys Val Ile Leu Asn Asp	
85 90 95	
tcc agc ccc caa gac caa gaa gag aga cct gca cag aga gac ttc tct	336
Ser Ser Pro Gln Asp Gln Glu Glu Arg Pro Ala Gln Arg Asp Phe Ser	
100 105 110	
ttc ccg agg ctc ccg agg gac cag ctc tac cgc cca cca tct aat gga	384
Phe Pro Arg Leu Pro Arg Asp Gln Leu Tyr Arg Pro Pro Ser Asn Gly	
115 120 125	
gtg gaa gcc cct gag gag tcc ttg gac ctt cct gca gag ctg aga cca	432
Val Glu Ala Pro Glu Glu Ser Leu Asp Leu Pro Ala Glu Leu Arg Pro	
130 135 140	
cat ggt ccc cag gct cca tcc cta gct gcc gtg cct aga ccc cct agc	480
His Gly Pro Gln Ala Pro Ser Leu Ala Ala Val Pro Arg Pro Pro Ser	
145 150 155 160	
aac ccc tta gcc cga gga act cta aga acc agc aat ttg cca gaa gaa	528
Asn Pro Leu Ala Arg Gly Thr Leu Arg Thr Ser Asn Leu Pro Glu Glu	
165 170 175	
tta cgg aaa gtc ttt atc act tat tct atg gac aca gcc atg gag gtg	576
Leu Arg Lys Val Phe Ile Thr Tyr Ser Met Asp Thr Ala Met Glu Val	
180 185 190	
gtg aaa ttt gtg aac ttt ctg ttg gtg aac ggc ttc caa act gcg att	624
Val Lys Phe Val Asn Phe Leu Leu Val Asn Gly Phe Gln Thr Ala Ile	
195 200 205	
gac ata ttt gag gat aga atc cgg ggt att gat atc att aaa tgg atg	672
Asp Ile Phe Glu Asp Arg Ile Arg Gly Ile Asp Ile Ile Lys Trp Met	
210 215 220	
gag cgc tat ctt cga gat aag aca gtg atg ata atc gta gca atc agc	720
Glu Arg Tyr Leu Arg Asp Lys Thr Val Met Ile Ile Val Ala Ile Ser	
225 230 235 240	
ccc aaa tac aaa cag gat gtg gaa ggc gct gag tcg cag ctg gac gag	768
Pro Lys Tyr Lys Gln Asp Val Glu Gly Ala Glu Ser Gln Leu Asp Glu	
245 250 255	
gac gag cat ggc tta cat act aag tac att cat cgg atg atg cag att	816
Asp Glu His Gly Leu His Thr Lys Tyr Ile His Arg Met Met Gln Ile	
260 265 270	
gag ttc ata agt cag gga agc atg aac ttc aga ttc atc cct gtg ctc	864
Glu Phe Ile Ser Gln Gly Ser Met Asn Phe Arg Phe Ile Pro Val Leu	
275 280 285	
ttc cca aat gcc aag aag gag cat gtg ccg acc tgg ctt cag aac act	912
Phe Pro Asn Ala Lys Lys Glu His Val Pro Thr Trp Leu Gln Asn Thr	
290 295 300	
cat gtt tac agc tgg ccc aag aat aag aaa aac atc ctg ctg cgg ctg	960
His Val Tyr Ser Trp Pro Lys Asn Lys Lys Asn Ile Leu Leu Arg Leu	
305 310 315 320	

ctc agg gag gaa gag tat gtg gct cct ccc cga ggc cct ctg ccc acc 1008  
 Leu Arg Glu Glu Glu Tyr Val Ala Pro Pro Arg Gly Pro Leu Pro Thr  
                     325                    330                    335

ctt cag gtg gta ccc ttg tgacgatggc cactccagct cagtgccagc 1056  
 Leu Gln Val Val Pro Leu  
                     340

ctgtttctcac agcattcttc tagcggagct ggctggtggc acccaggccc tggaacacct 1116

cttctacaga gtcctctgtc tcctgagtct gagttgtcct cgctgggctt ccagagcttc 1176

agtgcctgga tgctgcaggt gacagaaaca aacatctatg accacaaaaa ctctcatcac 1236

ttcagctact tttatgagtc ggtcagatgc tctgtgtcct tagaccagtc taaatcatgc 1296

tcaaataata aaatgattat tctttgt 1323

<210> 26

<211> 342

<212> PRT

<213> Unknown

<400> 26

Gln Asp Leu Pro Gly Pro Leu Arg Ser Arg Glu Leu Pro Pro Gln Phe  
     1                    5                    10                    15

Glu Leu Glu Arg Tyr Pro Met Asn Ala Gln Leu Leu Pro Pro His Pro  
                     20                    25                    30

Ser Pro Gln Ala Pro Trp Asn Cys Gln Tyr Tyr Cys Pro Gly Gly Pro  
                     35                    40                    45

Tyr His His Gln Val Pro His Gly His Gly Tyr Pro Pro Ala Ala Ala  
     50                    55                    60

Tyr Gln Gln Val Leu Gln Pro Ala Leu Pro Gly Gln Val Leu Pro Gly  
     65                    70                    75                    80

Ala Arg Ala Arg Gly Pro Arg Pro Val Gln Lys Val Ile Leu Asn Asp  
                     85                    90                    95

Ser Ser Pro Gln Asp Gln Glu Glu Arg Pro Ala Gln Arg Asp Phe Ser  
                     100                    105                    110

Phe Pro Arg Leu Pro Arg Asp Gln Leu Tyr Arg Pro Pro Ser Asn Gly  
                     115                    120                    125

Val Glu Ala Pro Glu Glu Ser Leu Asp Leu Pro Ala Glu Leu Arg Pro  
     130                    135                    140

His Gly Pro Gln Ala Pro Ser Leu Ala Ala Val Pro Arg Pro Pro Ser  
     145                    150                    155                    160

Asn Pro Leu Ala Arg Gly Thr Leu Arg Thr Ser Asn Leu Pro Glu Glu  
                     165                    170                    175

Leu Arg Lys Val Phe Ile Thr Tyr Ser Met Asp Thr Ala Met Glu Val

180					185					190					
Val	Lys	Phe	Val	Asn	Phe	Leu	Leu	Val	Asn	Gly	Phe	Gln	Thr	Ala	Ile
	195						200					205			
Asp	Ile	Phe	Glu	Asp	Arg	Ile	Arg	Gly	Ile	Asp	Ile	Ile	Lys	Trp	Met
	210					215					220				
Glu	Arg	Tyr	Leu	Arg	Asp	Lys	Thr	Val	Met	Ile	Ile	Val	Ala	Ile	Ser
	225					230					235				240
Pro	Lys	Tyr	Lys	Gln	Asp	Val	Glu	Gly	Ala	Glu	Ser	Gln	Leu	Asp	Glu
				245					250					255	
Asp	Glu	His	Gly	Leu	His	Thr	Lys	Tyr	Ile	His	Arg	Met	Met	Gln	Ile
		260					265							270	
Glu	Phe	Ile	Ser	Gln	Gly	Ser	Met	Asn	Phe	Arg	Phe	Ile	Pro	Val	Leu
		275					280					285			
Phe	Pro	Asn	Ala	Lys	Lys	Glu	His	Val	Pro	Thr	Trp	Leu	Gln	Asn	Thr
	290					295					300				
His	Val	Tyr	Ser	Trp	Pro	Lys	Asn	Lys	Lys	Asn	Ile	Leu	Leu	Arg	Leu
	305					310					315				320
Leu	Arg	Glu	Glu	Glu	Tyr	Val	Ala	Pro	Pro	Arg	Gly	Pro	Leu	Pro	Thr
			325					330						335	
Leu	Gln	Val	Val	Pro	Leu										
			340												

&lt;210&gt; 27

&lt;211&gt; 1026

&lt;212&gt; DNA

&lt;213&gt; reverse translation

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1)..(1026)

&lt;223&gt; n amy be a, c, g, or t

&lt;400&gt; 27

cargayytnc cnggnccnyt nmgnwsnmgn garytncnc cncarttyga rytngarmgn 60

tayccnatga aygcncaryt nytcccncn cayccnwsnc cncargcncc ntggaaytgy 120

cartaytayt gyccngngg nccntaycay caycargtnc cncayggncay yggntayccn 180

ccngcngcng cntaycarca rgtnytnear ccngcnytn cnggncargt nytncnggn 240

gcnmgngcnm gnggnccnmg nccngtnear aargtnathy tnaaygayws nwsnccncar 300

gaycargarg armgncngc ncarmgngay ttywsnttyc cnmgnytncc nmnggaycar 360

ytntaymgnc cncnwsnaa yggngtngar gcncngarg arwsnytna yytnccngcn 420

garytnmgnc cncayggnc ncarngcncc wsnytnngc cngtnccnmg nccnccnwsn 480

aayccnytn g cnmgnggnac nytnmgnacn wsnaayytnc cngargaryt nmgnaargtn 540  
ttyathacnt aywsnatgga yacngcnatg gargtngtna arttygtnaa ytttytnytn 600  
gtnaayggnt tyacaracngc nathgayath ttygargaym gnathmgngg nathgayath 660  
athaartgga tggarmgnta yytnmgngay aaracngtna tgathathgt ngcnathwsn 720  
ccnaartaya arcargaygt ngarggngcn garwsncary tngaygarga ygar cayggn 780  
ytncayacna artayathca ymgnatgatg carathgart tyathwsnca rggngwsnatg 840  
aayttymgnt tyathccngt nytnnttyccn aaygcnaara argarcaygt nccnacntgg 900  
ytncaraaya cncaygtnta ywsntggccn aaraayaara araayathyt nytnmgnytn 960  
ytmgngarg argartaygt ngcnccncn mgngngccny tncnacnyt ncargtngtn 1020  
ccnytn 1026

<210> 28

<211> 207

<212> PRT

<213> Unknown

<220>

<223> Description of Unknown Organism: primate; surmised  
Homo sapiens

<400> 28

Arg	Lys	Val	Trp	Ile	Ile	Tyr	Ser	Ala	Asp	His	Pro	Leu	Tyr	Val	Asp
1				5					10					15	
Val	Val	Leu	Lys	Phe	Ala	Gln	Phe	Leu	Leu	Thr	Ala	Cys	Gly	Thr	Glu
			20					25					30		
Val	Ala	Leu	Asp	Leu	Leu	Glu	Glu	Gln	Ala	Ile	Ser	Glu	Ala	Gly	Val
		35				40						45			
Met	Thr	Trp	Val	Gly	Arg	Gln	Lys	Gln	Glu	Met	Val	Glu	Ser	Asn	Ser
	50				55						60				
Lys	Ile	Ile	Val	Leu	Cys	Ser	Arg	Gly	Thr	Arg	Ala	Lys	Trp	Gln	Ala
65					70					75				80	
Leu	Leu	Gly	Arg	Gly	Ala	Pro	Val	Arg	Leu	Arg	Cys	Asp	His	Gly	Lys
			85						90					95	
Pro	Val	Gly	Asp	Leu	Phe	Thr	Ala	Ala	Met	Asn	Met	Ile	Leu	Pro	Asp
		100					105						110		
Phe	Lys	Arg	Pro	Ala	Cys	Phe	Gly	Thr	Tyr	Val	Val	Cys	Tyr	Phe	Ser
	115						120					125			
Glu	Val	Ser	Cys	Asp	Gly	Asp	Val	Pro	Asp	Leu	Phe	Gly	Ala	Ala	Pro
	130				135						140				
Arg	Tyr	Pro	Leu	Met	Asp	Arg	Phe	Glu	Glu	Val	Tyr	Phe	Arg	Ile	Gln
145				150					155					160	

```
<220>
<223> Description of Unknown Organism:rodent; surmised
      Mus musculus
```

<400>	29																
Arg	Lys	Val	Trp	Ile	Val	Tyr	Ser	Ala	Asp	His	Pro	Leu	Tyr	Val	Glu		
1				5					10					15			
Val	Val	Leu	Lys	Phe	Ala	Gln	Phe	Leu	Ile	Thr	Ala	Cys	Gly	Thr	Glu		
			20					25					30				
Val	Ala	Leu	Asp	Leu	Leu	Glu	Glu	Gln	Val	Ile	Ser	Glu	Val	Gly	Val		
		35					40					45					
Met	Thr	Trp	Val	Ser	Arg	Gln	Lys	Gln	Glu	Met	Val	Glu	Ser	Asn	Ser		
	50					55						60					
Lys	Ile	Ile	Ile	Leu	Cys	Ser	Arg	Gly	Thr	Gln	Ala	Lys	Trp	Lys	Ala		
65					70					75					80		
Ile	Leu	Gly	Trp	Ala	Glu	Pro	Ala	Val	Gln	Leu	Arg	Cys	Asp	His	Trp		
				85					90					95			
Lys	Pro	Ala	Gly	Asp	Leu	Phe	Thr	Ala	Ala	Met	Asn	Met	Ile	Leu	Pro		
			100					105					110				
Asp	Phe	Lys	Arg	Pro	Ala	Cys	Phe	Gly	Thr	Tyr	Val	Val	Cys	Tyr	Phe		
		115					120					125					
Ser	Gly	Ile	Cys	Ser	Glu	Arg	Asp	Val	Pro	Asp	Leu	Phe	Asn	Ile	Thr		
	130					135					140						
Ser	Arg	Tyr	Pro	Leu	Met	Asp	Arg	Phe	Glu	Glu	Val	Tyr	Phe	Arg	Ile		
145					150					155					160		
Gln	Asp	Leu	Glu	Met	Phe	Glu	Pro	Gly	Arg	Met	His	His	Val	Arg	Glu		
				165					170					175			
Leu	Thr	Gly	Asp	Asn	Tyr	Leu	Gln	Ser	Pro	Ser	Gly	Arg	Gln	Leu	Lys		
			180					185					190				
Glu	Ala	Val	Leu	Arg	Phe	Gln	Glu	Trp	Gln	Thr	Gln	Cys	Pro	Asp	Trp		
		195					200					205					



```
<210> 30
<211> 190
<212> PRT
<213> Unknown
```

**<220>**

<223> Description of Unknown Organism:worm; surmised  
Caenorabditis elegans

<400> 30

Val 1	Lys	Val	Met	Ile 5	Val	Tyr	Ala	Asp	Asp 10	Asn	Asp	Leu	His	Thr 15	Asp
Cys	Val	Lys	Lys 20	Leu	Val	Glu	Asn	Leu 25	Arg	Asn	Cys	Ala	Ser 30	Cys	Asp
Pro	Val	Phe 35	Asp	Leu	Glu	Lys	Leu 40	Ile	Thr	Ala	Glu	Ile 45	Val	Pro	Ser
Arg	Trp 50	Leu	Val	Asp	Gln	Ile 55	Ser	Ser	Leu	Lys	Lys 60	Phe	Ile	Ile	Val
Val 65	Ser	Asp	Cys	Ala	Glu 70	Lys	Ile	Leu	Asp 75	Thr	Glu	Ala	Ser	Glu 80	Thr
His	Gln	Leu	Val	Gln 85	Ala	Arg	Pro	Phe	Ala 90	Asp	Leu	Phe	Gly	Pro 95	Ala
Met	Glu	Met	Ile 100	Ile	Arg	Asp	Ala	Thr 105	His	Asn	Phe	Pro	Glu 110	Ala	Arg
Lys	Lys	Tyr 115	Ala	Val	Val	Arg	Phe 120	Asn	Tyr	Ser	Pro	His 125	Val	Pro	Pro
Asn	Leu	Ala	Ile 130	Leu	Asn	Leu 135	Pro	Thr	Phe	Ile	Pro 140	Glu	Gln	Phe	Ala
Gln 145	Leu	Thr	Ala	Phe	Leu 150	His	Asn	Val	Glu	His 155	Thr	Glu	Arg	Ala	Asn 160
Val	Thr	Gln	Asn 165	Ile	Ser	Glu	Ala	Gln	Ile 170	His	Glu	Trp	Asn 175	Leu	Cys
Ala	Ser	Arg	Met 180	Met	Ser	Phe	Phe	Val 185	Arg	Asn	Pro	Asn 190	Trp		

<210> 31  
<211> 178  
<212> PRT  
<213> Unknown

**<220>**

<223> Description of Unknown Organism:worm; surmised  
Caenorabditis elegans

&lt;400&gt; 31

```

Phe Lys Val Met Leu Val Cys Pro Glu Val Ser Gly Arg Asp Glu Asp
 1             5             10             15

Phe Met Met Arg Ile Ala Asp Ala Leu Lys Lys Ser Asn Asn Lys Val
      20             25             30

Val Cys Asp Arg Trp Phe Glu Asp Ser Lys Asn Ala Glu Glu Asn Met
      35             40             45

Leu His Trp Val Tyr Glu Gln Thr Lys Ile Ala Glu Lys Ile Ile Val
 50             55             60

Phe His Ser Ala Tyr Tyr His Pro Arg Cys Gly Ile Tyr Asp Val Ile
 65             70             75             80

Asn Asn Phe Phe Pro Cys Thr Asp Pro Arg Leu Ala His Ile Ala Leu
      85             90             95

Thr Pro Glu Ala Gln Arg Ser Val Pro Lys Glu Val Glu Tyr Val Leu
      100             105             110

Pro Arg Asp Gln Lys Leu Leu Glu Asp Ala Phe Asp Ile Thr Ile Ala
      115             120             125

Asp Pro Leu Val Ile Asp Ile Pro Ile Glu Asp Val Ala Ile Pro Glu
      130             135             140

Asn Val Pro Ile His His Glu Ser Cys Asp Ser Ile Asp Ser Arg Asn
      145             150             155             160

Asn Ser Lys Thr His Ser Thr Asp Ser Gly Val Ser Ser Leu Ser Ser
      165             170             175

Asn Ser

```

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
29 November 2001 (29.11.2001)

PCT

(10) International Publication Number  
**WO 01/090358 A3**

(51) International Patent Classification<sup>7</sup>: **C12N 15/12**,  
C07K 14/715, 16/18, G01N 33/53, C12N 5/10

(21) International Application Number: PCT/US01/16767

(22) International Filing Date: 23 May 2001 (23.05.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/206,862 24 May 2000 (24.05.2000) US

(71) Applicant: **SCHERING CORPORATION** [US/US];  
2000 Galloping Hill Road, Kenilworth, NJ 07033-0530  
(US).

(72) Inventor: **GORMAN, Daniel, M.**; 6371 Central Avenue,  
Newark, CA 94560 (US).

(74) Agent: **ZARADIC, Sandy**; Schering-Plough Corpora-  
tion, Patent Department, K-6-1, 1990, 2000 Galloping Hill  
Road, Kenilworth, NJ 07033-0530 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CZ,

DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, HR, HU,  
ID, IL, IN, IS, JP, KG, KR, KZ, LC, LK, LR, LT, LU, LV,  
MA, MD, MG, MK, MN, MX, MZ, NO, NZ, PL, PT, RO,  
RU, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UZ, VN,  
YU, ZA.

(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Declaration under Rule 4.17:**

*as to the applicant's entitlement to claim the priority of the  
earlier application (Rule 4.17(iii)) for all designations*

**Published:**

— *with international search report*

(88) Date of publication of the international search report:

23 January 2003

*For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*

WO 01/090358 A3

(54) Title: MAMMALIAN RECEPTOR PROTEINS; RELATED REAGENTS AND METHODS

(57) Abstract: Nucleic acids encoding mammalian, e.g., primate, receptors, purified receptor proteins and fragments thereof. Anti-  
bodies, both polyclonal and monoclonal, are also provided. Methods of using the compositions for both diagnostic and therapeutic  
utilities are described.

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/16767

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C07K14/715 C07K16/18 G01N33/53 C12N5/10

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

SEQUENCE SEARCH, EMBL, EPO-Internal, MEDLINE, BIOSIS, WPI Data, PAJ, CHEM ABS Data, SCISEARCH, EMBASE

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96 29408 A (IMMUNEX CORP) 26 September 1996 (1996-09-26) page 2, line 35 -page 15, line 4 ---	1-18
X	YAO Z ET AL: "MOLECULAR CHARACTERIZATION OF THE HUMAN INTERLEUKIN (IL)-17 RECEPTOR" CYTOKINE, ACADEMIC PRESS LTD, PHILADELPHIA, PA, US, vol. 9, no. 11, November 1997 (1997-11), pages 794-800, XP000867704 ISSN: 1043-4666 page 795; figure 2 --- -/--	1-4, 6, 12-15



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

## \* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "I" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

12 August 2002

Date of mailing of the international search report

29.08.02

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3018

Authorized officer

Steffen, P

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/16767

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EMBL 'Online!  EBI; 18 February 2000 (2000-02-18)  BLOECKER ET AL.: "Homo sapiens mRNA; cDNA  DKFZp434N1928"  Database accession no. AL133097  XP002183453  the whole document</p> <p>---</p>	1-4, 6, 12-15
A	<p>WO 99 14240 A (HUMAN GENOME SCIENCES INC  ; RUBEN STEVEN M (US); SHI YANGGU (US))  25 March 1999 (1999-03-25)  the whole document</p> <p>---</p>	
A	<p>TIAN E ET AL: "EVI27 ENCODES A NOVEL  MEMBRANE PROTEIN WITH HOMOLOGY TO THE IL17  RECEPOR"  ONCOGENE, BASINGSTOKE, HANTS, GB,  vol. 19, no. 17,  20 April 2000 (2000-04-20), pages  2098-2109, XP008000240  ISSN: 0950-9232  the whole document</p> <p>---</p>	
A	<p>SHI YANGGU ET AL: "A novel cytokine  receptor-ligand pair: Identification,  molecular characterization, and in vivo  immunomodulatory activity."  JOURNAL OF BIOLOGICAL CHEMISTRY (JBC  PAPERS IN PRESS, DOI  10.1074/JBC.M910228199),  vol. 275, no. 25,  3 April 2000 (2000-04-03), pages  19167-19176, XP002197927  ISSN: 0021-9258  the whole document</p> <p>---</p>	
A	<p>FOSSIEZ F ET AL: "INTERLEUKIN-17"  INTERNATIONAL REVIEWS OF IMMUNOLOGY,  HARWOOD ACADEMIC PUBLISHERS, LONDON, GB,  vol. 16, no. 5/6, 1998, pages 541-551,  XP000867763  ISSN: 0883-0185  the whole document</p> <p>---</p>	
E	<p>WO 01 68859 A (AMGEN INC ; JING SHUQIAN  (US)) 20 September 2001 (2001-09-20)  page 2, line 19 -page 10, line 27;  examples 1-4</p> <p>---</p>	1-18
E	<p>WO 01 46420 A (GENENTECH INC)  28 June 2001 (2001-06-28)  page 5, line 1 -page 16, line 17; figures  17, 18</p> <p>---</p>	1-18

-/-

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/16767

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 55865 A, (GENESIS RES & DEV CORP LTD) 4 November 1999 (1999-11-04) SEQ ID NO's 125 and 303 and corresponding cDNA's page 3 -page 17 ----	1-18
X	DATABASE EMBL 'Online! EBI; 22 July 1999 (1999-07-22) NCI-CGAP: "ty30c03.x1 NCI_CGAP_UT2 Homo sapiens cDNA clone IMAGE:2280580 3' mRNA sequence" Database accession no. AI861981 XP002209553 the whole document ----	12-18
X	DATABASE EMBL 'Online! EBI; 21 October 1999 (1999-10-21) MARRA ET AL.: "u191g04.y1 Sugano mouse kidney mkia Mus musculus cDNA clone IMAGE:2159478 5', mRNA sequence" Database accession no. AW107583 XP002209554 the whole document -----	12-18

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US 01/16767

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claims Nos.: 19, 20  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:  
see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☒ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☒ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-18 (all partly)

Compositions comprising primate DCRS8 polypeptides and nucleic acid sequences (SEQ ID NO's 14 and 13, respectively) as well as further embodiments relating to the said polypeptides and nucleic acid sequences.

2. Claims: 1-18 (all partly)

Compositions comprising primate or rodent DCRS9 polypeptides and nucleic acid sequences (SEQ ID NO's 16, 19 and 17, 20, respectively) as well as further embodiments relating to the said polypeptides and nucleic acid sequences.



## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 19, 20

Present claims 19 and 20 relate to a method defined by reference to a desirable characteristic or property, namely contacting a cell with an unspecified agonist or antagonist of a mammalian protein of the application (e.g. DCRS8 or DCRS9).

The claims cover all methods having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such methods. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the method by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, no search has been carried out for claims 19 and 20.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/16767

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9629408	A	26-09-1996	AU 696775 B2	17-09-1998
			AU 5526396 A	08-10-1996
			CA 2215394 A1	26-09-1996
			EP 0817847 A1	14-01-1998
			JP 11503309 T	26-03-1999
			NO 974258 A	24-11-1997
			NZ 306653 A	29-03-1999
			US 6096305 A	01-08-2000
			US 6072033 A	06-06-2000
			US 6197525 B1	06-03-2001
			US 6191104 B1	20-02-2001
			US 6100235 A	08-08-2000
			US 6072037 A	06-06-2000
			WO 9629408 A1	26-09-1996
			US 5869286 A	09-02-1999
WO 9914240	A	25-03-1999	AU 9482498 A	05-04-1999
			EP 1015488 A1	05-07-2000
			WO 9914240 A1	25-03-1999
WO 0168859	A	20-09-2001	AU 4754501 A	24-09-2001
			WO 0168859 A2	20-09-2001
			US 2002045213 A1	18-04-2002
WO 0146420	A	28-06-2001	AU 2590901 A	03-07-2001
			WO 0146420 A2	28-06-2001
			AU 2055401 A	12-06-2001
			AU 2224800 A	28-09-2000
			AU 2596700 A	28-09-2000
			AU 2877900 A	28-09-2000
			AU 3514400 A	28-09-2000
			EP 1220905 A2	10-07-2002
			WO 0053754 A1	14-09-2000
			WO 0053756 A2	14-09-2000
			WO 0053758 A2	14-09-2000
			WO 0140466 A2	07-06-2001
			US 2002094416 A1	18-07-2002
			AU 3434601 A	12-06-2001
			AU 7573000 A	26-03-2001
			EP 1208202 A2	29-05-2002
			WO 0116318 A2	08-03-2001
			WO 0140465 A2	07-06-2001
			AU 2883600 A	28-09-2000
			AU 3381600 A	28-09-2000
			AU 3774300 A	18-12-2000
			AU 5152700 A	02-01-2001
			AU 5459900 A	18-12-2000
			AU 6391000 A	19-02-2001
			EP 1159419 A1	05-12-2001
			EP 1210418 A1	05-06-2002
			EP 1208195 A2	29-05-2002
			EP 1212417 A2	12-06-2002
			EP 1185648 A2	13-03-2002
			WO 0053757 A2	14-09-2000
			WO 0073454 A1	07-12-2000
			WO 0073445 A2	07-12-2000
			WO 0077037 A2	21-12-2000
			WO 0073348 A2	07-12-2000

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/16767

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0146420	A	WO 0073452 A2	07-12-2000
		WO 0109327 A2	08-02-2001
		US 2002072496 A1	13-06-2002
		US 2002072067 A1	13-06-2002
		US 2002072092 A1	13-06-2002
		US 2002072497 A1	13-06-2002
		AU 1916701 A	16-07-2001
		AU 6802801 A	24-09-2001
		WO 0149715 A2	12-07-2001
		WO 0168848 A2	20-09-2001
		US 2002098505 A1	25-07-2002
		US 2002098506 A1	25-07-2002
		US 2002098507 A1	25-07-2002
		US 2002090681 A1	11-07-2002
		US 2002010137 A1	24-01-2002
		US 2002058309 A1	16-05-2002
WO 9955865	A	04-11-1999	
		US 6150502 A	21-11-2000
		AU 3961499 A	16-11-1999
		CA 2327317 A1	04-11-1999
		CN 1307637 T	08-08-2001
		EP 1073740 A1	07-02-2001
		JP 2002512798 T	08-05-2002
		WO 9955865 A1	04-11-1999
		NZ 507728 A	28-06-2002

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
3 January 2002 (03.01.2002)

PCT

(10) International Publication Number  
**WO 02/00690 A2**

- (51) International Patent Classification<sup>7</sup>: **C07K 14/00**
- (21) International Application Number: PCT/US01/19692
- (22) International Filing Date: 20 June 2001 (20.06.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
- |                |                                |    |
|----------------|--------------------------------|----|
| 60/213,637     | 23 June 2000 (23.06.2000)      | US |
| 60/219,556     | 20 July 2000 (20.07.2000)      | US |
| 60/220,624     | 25 July 2000 (25.07.2000)      | US |
| 60/220,664     | 25 July 2000 (25.07.2000)      | US |
| PCT/US00/20710 | 28 July 2000 (28.07.2000)      | US |
| 60/222,695     | 2 August 2000 (02.08.2000)     | US |
| 09/643,657     | 17 August 2000 (17.08.2000)    | US |
| PCT/US00/23522 | 23 August 2000 (23.08.2000)    | US |
| PCT/US00/23328 | 24 August 2000 (24.08.2000)    | US |
| 60/230,978     | 7 September 2000 (07.09.2000)  | US |
| 60/000,000     | 15 September 2000 (15.09.2000) | US |
| 09/664,610     | 18 September 2000 (18.09.2000) | US |
| 09/665,350     | 18 September 2000 (18.09.2000) | US |
| 60/242,922     | 24 October 2000 (24.10.2000)   | US |
| 09/709,238     | 8 November 2000 (08.11.2000)   | US |
| PCT/US00/30952 | 8 November 2000 (08.11.2000)   | US |
| PCT/US00/30873 | 10 November 2000 (10.11.2000)  | US |
| PCT/US00/32678 | 1 December 2000 (01.12.2000)   | US |
| 09/747,259     | 20 December 2000 (20.12.2000)  | US |
| PCT/US00/34956 | 20 December 2000 (20.12.2000)  | US |
| 09/767,609     | 22 January 2001 (22.01.2001)   | US |
| 09/796,498     | 28 February 2001 (28.02.2001)  | US |
| PCT/US01/06520 | 28 February 2001 (28.02.2001)  | US |
| PCT/US01/06666 | 1 March 2001 (01.03.2001)      | US |
| 09/802,706     | 9 March 2001 (09.03.2001)      | US |
| 09/808,689     | 14 March 2001 (14.03.2001)     | US |
| 09/816,744     | 22 March 2001 (22.03.2001)     | US |
| 09/828,366     | 5 April 2001 (05.04.2001)      | US |
- (71) Applicant (for all designated States except US): **GENENTECH, INC.** [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **BAKER, Kevin, P.** [GB/US]; 14006 Indian Run Drive, Darnestown, MD 20878 (US). **FERRARA, Napoleone** [US/US]; #704, 2090 Pacific Avenue, San Francisco, CA 94109 (US). **GERBER, Hanspeter** [CH/US]; #5, 1121 Tennessee Street, San Francisco, CA 94107 (US). **GERRITSEN, Mary, E.** [CA/US]; 541 Parrott Drive, San Mateo, CA 94402 (US). **GODDARD, Audrey** [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). **GODOWSKI, Paul, J.** [US/US]; 2627 Easton Drive, Burlingame, CA 94010 (US). **GURNEY, Austin, L.** [US/US]; 1 Debbie Lane, Belmont, CA 94002 (US). **HILLAN, Kenneth, J.** [GB/US]; 64 Seward Street, San Francisco, CA 94114 (US). **MARSTERS, Scot, A.** [US/US]; 990 Cherry Street, San Carlos, CA 94070 (US). **PAN, James** [CA/US]; 2705 Coronet Boulevard, Belmont, CA 94002 (US). **PAONI, Nicholas, F.** [US/US]; 1756 Terrace Drive, Belmont, CA 94002 (US). **STEPHAN, Jean-Philippe, F.** [FR/US]; 320 C Lansdale Avenue, Millbrae, CA 94030 (US). **WATANABE, Colin, K.** [US/US]; 128 Corliss Drive, Moraga, CA 94556 (US). **WILLIAMS, P., Mickey** [US/US]; 509 Alto Avenue, Half Moon Bay, CA 94019 (US). **WOOD, William, I.** [US/US]; 35 Southdown Court, Hillsborough, CA 94010 (US). **YE, Weilan** [CN/US]; 119 Barkentine Street, Foster City, CA 94404 (US).
- (74) Agents: **AGARWAL, Atulya, R.** et al.; c/o GENENTECH, INC., MS49, 1 DNA Way, South San Francisco, CA 94080-4990 (US).

[Continued on next page]

WO 02/00690 A2

(54) Title: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS

(57) Abstract: Compositions and methods are disclosed for stimulating or inhibiting angiogenesis and/or cardiovascularization in mammals, including humans. Pharmaceutical compositions are based on polypeptides or antagonists thereto that have been identified for one or more of these uses. Disorders that can be diagnosed, prevented, or treated by the compositions herein include trauma such as wounds, various cancers, and disorders of the vessels including atherosclerosis and cardiac hypertrophy. In addition, the present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.



(81) **Designated States (national):** AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— without international search report and to be republished upon receipt of that report

(84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

**COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF DISORDERS  
INVOLVING ANGIOGENESIS**

5        1.        Field of the Invention

The present invention relates to compositions and methods useful for the modulation (e.g., promotion or inhibition) of angiogenesis and/or cardiovascularization in mammals in need of such biological effect. The present invention further relates to the diagnosis and treatment of disorders involving angiogenesis (e.g., cardiovascular as well as oncological disorders).

10

2.        Background of the Invention

2.1.      Angiogenesis

Angiogenesis, defined as the growth or sprouting of new blood vessels from existing vessels, is a complex process that primarily occurs during embryonic development. Under normal physiological conditions in adults, angiogenesis takes place only in very restricted situations such as hair growth and wounding healing (Auerbach, W. and Auerbach, R., 1994, *Pharmacol Ther* 63(3):265-3 11; Ribatti et al., 1991, *Haematologica* 76(4):3 11-20; Risau, 1997, *Nature* 386(6626):67 1-4). Unregulated angiogenesis has gradually been recognized to be responsible for a wide range of disorders, including, but not limited to cardiovascular disease, cancer, rheumatoid arthritis, psoriasis and diabetic retinopathy (Folkman, 1995, *Nat Med* 1(1):27-31; Isner, 1999, *Circulation* 99(13): 1653-5; Koch, 1998, *Arthritis Rheum* 41(6):951-62; Walsh, 1999, *Rheumatology (Oxford)* 38(2):103-12; Ware and Simons, 1997, *Nat Med* 3(2): 158-64).

20

2.2.      Cardiac Disorders and Factors

Heart failure affects approximately five million Americans, and new cases of heart failure number about 400,000 each year. It is the single most frequent cause of hospitalization for people age 65 and older in the United States. Recent advances in the management of acute cardiac diseases, including acute myocardial infarction, are resulting in an expanding patient population that will eventually develop chronic heart failure. From 1979 to 1995, hospitalizations for congestive heart failure (CHF) rose from 377,000 to 872,000 (a 130 percent increase) and CHF deaths increased 116 percent.

25

CHF is a syndrome characterized by left ventricular dysfunction, reduced exercise tolerance, impaired quality of life, and markedly shortened life expectancy. The sine qua non of heart failure is an inability of the heart to pump blood at a rate sufficient to meet the metabolic needs of the body's tissues (in other words, there is insufficient cardiac output).

30

At least four major compensatory mechanisms are activated in the setting of heart failure to boost cardiac output, including peripheral vasoconstriction, increased heart rate, increased cardiac contractility, and increased plasma volume. These effects are mediated primarily by the sympathetic nervous system and the renin-angiotensin system. See, Eichhorn, American Journal of Medicine, 104: 163-169 (1998). Increased output from the

35

sympathetic nervous system increases vascular tone, heart rate, and contractility. Angiotensin II elevates blood pressure by 1) directly stimulating vascular smooth muscle contraction, 2) promoting plasma volume expansion by stimulating aldosterone and antidiuretic hormone secretion, 3) stimulating sympathetic-mediated vascular tone, and 4) catalyzing the degradation of bradykinin, which has vasodilatory and natriuretic activity. *See*, review by Brown and Vaughan, Circulation, 97: 1411-1420 (1998). As noted below, angiotensin II may also have directly deleterious effects on the heart by promoting myocyte necrosis (impairing systolic function) and intracardiac fibrosis (impairing diastolic and in some cases systolic function). *See*, Weber, Circulation, 96: 4065-4082 (1998).

A consistent feature of congestive heart failure (CHF) is cardiac hypertrophy, an enlargement of the heart that is activated by both mechanical and hormonal stimuli and enables the heart to adapt to demands for increased cardiac output. Morgan and Baker, Circulation, 83: 13-25 (1991). This hypertrophic response is frequently associated with a variety of distinct pathological conditions such as hypertension, aortic stenosis, myocardial infarction, cardiomyopathy, valvular regurgitation, and intracardiac shunt, all of which result in chronic hemodynamic overload.

Hypertrophy is generally defined as an increase in size of an organ or structure independent of natural growth that does not involve tumor formation. Hypertrophy of the heart is due either to an increase in the mass of the individual cells (myocytes), or to an increase in the number of cells making up the tissue (hyperplasia), or both. While the enlargement of an embryonic heart is largely dependent on an increase in myocyte number (which continues until shortly after birth), post-natal cardiac myocytes lose their proliferative capacity. Further growth occurs through hypertrophy of the individual cells.

Adult myocyte hypertrophy is initially beneficial as a short term response to impaired cardiac function by permitting a decrease in the load on individual muscle fibers. With severe, long-standing overload, however, the hypertrophied cells begin to deteriorate and die. Katz, "Heart Failure", in: Katz A.M. ed., Physiology of the Heart (New York: Raven Press, 1992) pp. 638-668. Cardiac hypertrophy is a significant risk factor for both mortality and morbidity in the clinical course of heart failure. Katz, Trends Cardiovasc. Med., 5: 37-44 (1995). For further details of the causes and pathology of cardiac hypertrophy *see, e.g.*, Heart Disease, A Textbook of Cardiovascular Medicine, Braunwald, E. ed. (W.B. Saunders Co., 1988), Chapter 14, "Pathophysiology of Heart Failure."

On a cellular level, the heart is composed of myocytes and surrounding support cells, generically called non-myocytes. While non-myocytes are primarily fibroblast/mesenchymal cells, they also include endothelial and smooth muscle cells. Indeed, although myocytes make up most of the adult myocardial mass, they represent only about 30% of the total cell numbers present in heart. In response to hormonal, physiological, hemodynamic, and pathological stimuli, adult ventricular muscle cells can adapt to increased workloads through the activation of a hypertrophic process. This response is characterized by an increase in myocyte cell size and contractile protein content of individual cardiac muscle cells, without concomitant cell division and activation of embryonic genes, including the gene for atrial natriuretic peptide (ANP). Chien *et al.*, FASEB J., 5: 3037-3046 (1991); Chien *et al.*, Annu. Rev. Physiol., 55: 77-95 (1993). An increment in myocardial mass as a result of an increase in myocyte size that is associated with an accumulation of interstitial collagen within the extracellular matrix and around intramyocardial coronary arteries has been described in left ventricular hypertrophy secondary to pressure overload

in humans. Caspari *et al.*, Cardiovasc. Res., 11: 554-558 (1977); Schwarz *et al.*, Am. J. Cardiol., 42: 895-903 (1978); Hess *et al.*, Circulation, 63: 360-371 (1981); Pearlman *et al.*, Lab. Invest., 46: 158-164 (1982).

It has also been suggested that paracrine factors produced by non-myocyte supporting cells may additionally be involved in the development of cardiac hypertrophy, and various non-myocyte derived hypertrophic factors, such as, leukocyte inhibitory factor (LIF) and endothelin, have been identified. Metcalf, Growth Factors, 7: 169-173 (1992); Kurzrock *et al.*, Endocrine Reviews, 12: 208-217 (1991); Inoue *et al.*, Proc. Natl. Acad. Sci. USA, 86: 2863-2867 (1989); Yanagisawa and Masaki, Trends Pharm. Sci., 10: 374-378 (1989); U.S. Patent No. 5,573,762 (issued November 12, 1996). Further exemplary factors that have been identified as potential mediators of cardiac hypertrophy include cardiotrophin-1 (CT-1) (Pennica *et al.*, Proc. Nat. Acad. Sci. USA, 92: 1142-1146 (1995)), catecholamines, adrenocorticosteroids, angiotensin, and prostaglandins.

At present, the treatment of cardiac hypertrophy varies depending on the underlying cardiac disease. Catecholamines, adrenocorticosteroids, angiotensin, prostaglandins, LIF, endothelin (including endothelin-1, -2, and -3 and big endothelin), and CT-1 are among the factors identified as potential mediators of hypertrophy. For example, beta-adrenergic receptor blocking drugs (beta-blockers, *e.g.*, propranolol, timolol, tertalolol, carteolol, nadolol, betaxolol, penbutolol, acetobutolol, atenolol, metoprolol, carvedilol, etc.) and verapamil have been used extensively in the treatment of hypertrophic cardiomyopathy. The beneficial effects of beta-blockers on symptoms (*e.g.*, chest pain) and exercise tolerance are largely due to a decrease in the heart rate with a consequent prolongation of diastole and increased passive ventricular filling. Thompson *et al.*, Br. Heart J., 44: 488-98 (1980); Harrison *et al.*, Circulation, 29: 84-98 (1964). Verapamil has been described to improve ventricular filling and probably reducing myocardial ischemia. Bonow *et al.*, Circulation, 72: 853-64 (1985).

Nifedipine and diltiazem have also been used occasionally in the treatment of hypertrophic cardiomyopathy. Lorell *et al.*, Circulation, 65: 499-507 (1982); Betocchi *et al.*, Am. J. Cardiol., 78: 451-457 (1996). However, because of its potent vasodilating properties, nifedipine may be harmful, especially in patients with outflow obstruction. Disopyramide has been used to relieve symptoms by virtue of its negative inotropic properties. Pollick, N. Engl. J. Med., 307: 997-999 (1982). In many patients, however, the initial benefits decrease with time. Wigle *et al.*, Circulation, 92: 1680-1692 (1995). Antihypertensive drug therapy has been reported to have beneficial effects on cardiac hypertrophy associated with elevated blood pressure. Examples of drugs used in antihypertensive therapy, alone or in combination, are calcium antagonists, *e.g.*, nitrendipine; adrenergic receptor blocking agents, *e.g.*, those listed above; angiotensin converting enzyme (ACE) inhibitors such as quinapril, captopril, enalapril, ramipril, benazepril, fosinopril, and lisinopril; diuretics, *e.g.*, chlorothiazide, hydrochlorothiazide, hydroflumethazide, methylchlorothiazide, benzthiazide, dichlorphenamide, acetazolamide, and indapamide; and calcium channel blockers, *e.g.*, diltiazem, nifedipine, verapamil, and nicardipine.

For example, treatment of hypertension with diltiazem and captopril showed a decrease in left ventricular muscle mass, but the Doppler indices of diastolic function did not normalize. Szlachet *et al.*, Am. J. Cardiol., 63: 198-201 (1989); Shahi *et al.*, Lancet, 336: 458-461 (1990). These findings were interpreted to indicate that excessive amounts of interstitial collagen may remain after regression of left ventricular hypertrophy. Rossi *et al.*, Am. Heart J., 124: 700-709 (1992). Rossi *et al.*, *supra*, investigated the effect of captopril on the prevention and



regression of myocardial cell hypertrophy and interstitial fibrosis in pressure overload cardiac hypertrophy, in experimental rats.

Agents that increase cardiac contractility directly (ionotropic agents) were initially thought to benefit patients with heart failure because they improved cardiac output in the short term. However, all positive inotropic agents except digoxigenin have been found to result in increased long-term mortality, in spite of short-term improvements in cardiac performance. Massie, Curr. Op. in Cardiology, 12: 209-217 (1997); Reddy *et al.*, Curr. Opin. Cardiol., 12: 233-241 (1997). Beta-adrenergic receptor blockers have recently been advocated for use in heart failure. Evidence from clinical trials suggests that improvements in cardiac function can be achieved without increased mortality, though documented improvements of patient survival have not yet been demonstrated. See also, U.S. Pat. Nos. 5,935,924, 5,624,806; 5,661,122; and 5,610,134 and WO 95/28173 regarding the use of cardiotropin-1 or antagonists thereof, or growth hormone and/or insulin-like growth factor-I in the treatment of CHF. Another treatment modality is heart transplantation, but this is limited by the availability of donor hearts.

Endothelin is a vasoconstricting peptide comprising 21 amino acids, isolated from swine arterial endothelial culture supernatant and structurally determined. Yanagisawa *et al.*, Nature, 332: 411-415 (1988). Endothelin was later found to exhibit various actions, and endothelin antibodies as endothelin antagonists have proven effective in the treatment of myocardial infarction, renal failure, and other diseases. Since endothelin is present in live bodies and exhibits vasoconstricting action, it is expected to be an endogenous factor involved in the regulation of the circulatory system, and may be associated with hypertension, cardiovascular diseases such as myocardial infarction, and renal diseases such as acute renal failure. Endothelin antagonists are described, for example, in U.S. Pat. No. 5,773,414; JP Pat. Publ. 3130299/1991, EP 457,195; EP 460,679; and EP 552,489. A new endothelin B receptor for identifying endothelin receptor antagonists is described in U.S. Pat. No. 5,773,223.

Current therapy for heart failure is primarily directed to using angiotensin-converting enzyme (ACE) inhibitors, such as captopril, and diuretics. These drugs improve hemodynamic profile and exercise tolerance and reduce the incidence of morbidity and mortality in patients with CHF. Kramer *et al.*, Circulation, 67(4): 807-816 (1983); Captopril Multicenter Research Group, J.A.C.C., 2(4): 755-763 (1983); The CONSENSUS Trial Study Group, N. Engl. J. Med., 316(23): 1429-1435 (1987); The SOLVD Investigators, N. Engl. J. Med., 325(5): 293-302 (1991). Further, they are useful in treating hypertension, left ventricular dysfunction, atherosclerotic vascular disease, and diabetic nephropathy. Brown and Vaughan, *supra*. However, despite proven efficacy, response to ACE inhibitors has been limited. For example, while prolonging survival in the setting of heart failure, ACE inhibitors appear to slow the progression towards end-stage heart failure, and substantial numbers of patients on ACE inhibitors have functional class III heart failure.

Moreover, improvement of functional capacity and exercise time is only small and mortality, although reduced, continues to be high. The CONSENSUS Trial Study Group, N. Engl. J. Med., 316(23): 1429-1435 (1987); The SOLVD Investigators, N. Engl. J. Med., 325(5): 293-302 (1991); Cohn *et al.*, N. Engl. J. Med., 325(5): 303-310 (1991); The Captopril-Digoxin Multicenter Research Group, JAMA, 259(4): 539-544 (1988). Hence, ACE inhibitors consistently appear unable to relieve symptoms in more than 60% of heart failure patients and reduce mortality of heart failure only by approximately 15-20%. For further adverse effects, see Brown and Vaughan,

*supra*.

An alternative to ACE inhibitors is represented by specific AT1 receptor antagonists. Clinical studies are planned to compare the efficacy of these two modalities in the treatment of cardiovascular and renal disease. However, animal model data suggests that the ACE/Ang II pathway, while clearly involved in cardiac hypertrophy, is not the only, or even the primary pathway active in this role. Mouse genetic "knockout" models have been made to test individual components of the pathway. In one such model, the primary cardiac receptor for Ang II, AT sub 1A, has been genetically deleted; these mice do not develop hypertrophy when Ang II is given experimentally (confirming the basic success of the model in eliminating hypertrophy secondary to Ang II). However, when the aorta is constricted in these animals (a model of hypertensive cardiac stress), the hearts still become hypertrophic. This suggests that alternative signaling pathways, not depending on this receptor (AT sub 1A), are activated in hypertension. ACE inhibitors would presumably not be able to inhibit these pathways. See, Harada *et al.*, Circulation, 97: 1952-1959 (1998). See also, Homcy, Circulation, 97: 1890-1892 (1998) regarding the enigma associated with the process and mechanism of cardiac hypertrophy.

About 750,000 patients suffer from acute myocardial infarction (AMI) annually, and approximately one-fourth of all deaths in the United States are due to AMI. In recent years, thrombolytic agents, *e.g.*, streptokinase, urokinase, and in particular tissue plasminogen activator (t-PA) have significantly increased the survival of patients who suffered myocardial infarction. When administered as a continuous intravenous infusion over 1.5 to 4 hours, t-PA produces coronary patency at 90 minutes in 69% to 90% of the treated patients. Topol *et al.*, Am. J. Cardiol., 61: 723-728 (1988); Neuhaus *et al.*, J. Am. Coll. Cardiol., 12: 581-587 (1988); Neuhaus *et al.*, J. Am. Coll. Cardiol., 14: 1566-1569 (1989). The highest patency rates have been reported with high dose or accelerated dosing regimens. Topol, J. Am. Coll. Cardiol., 15: 922-924 (1990). t-PA may also be administered as a single bolus, although due to its relatively short half-life, it is better suited for infusion therapy. Tebbe *et al.*, Am. J. Cardiol., 64: 448-453 (1989). A t-PA variant, specifically designed to have longer half-life and very high fibrin specificity, TNK t-PA (a T103N, N117Q, KHRH(296-299)AAAA t-PA variant, Keyt *et al.*, Proc. Natl. Acad. Sci. USA, 91: 3670-3674 (1994)) is particularly suitable for bolus administration. However, despite all these advances, the long-term prognosis of patient survival depends greatly on the post-infarction monitoring and treatment of the patients, which should include monitoring and treatment of cardiac hypertrophy.

### 2.3. Growth Factors

Various naturally occurring polypeptides reportedly induce the proliferation of endothelial cells. Among those polypeptides are the basic and acidic fibroblast growth factors (FGF) (Burgess and Maciag, Annual Rev. Biochem., 58: 575 (1989)), platelet-derived endothelial cell growth factor (PD-ECGF) (Ishikawa *et al.*, Nature, 338: 557 (1989)), and vascular endothelial growth factor (VEGF). Leung *et al.*, Science, 246: 1306 (1989); Ferrara and Henzel, Biochem. Biophys. Res. Commun., 161: 851 (1989); Tischer *et al.*, Biochem. Biophys. Res. Commun., 165: 1198 (1989); EP 471,754B granted July 31, 1996.

Media conditioned by cells transfected with the human VEGF (hVEGF) cDNA promoted the proliferation of capillary endothelial cells, whereas control cells did not. Leung *et al.*, Science, 246: 1306 (1989). Several

additional cDNAs were identified in human cDNA libraries that encode 121-, 189-, and 206-amino acid isoforms of hVEGF (also collectively referred to as hVEGF-related proteins). The 121-amino acid protein differs from hVEGF by virtue of the deletion of the 44 amino acids between residues 116 and 159 in hVEGF. The 189-amino acid protein differs from hVEGF by virtue of the insertion of 24 amino acids at residue 116 in hVEGF, and apparently is identical to human vascular permeability factor (hVPF). The 206-amino acid protein differs from hVEGF by virtue of an insertion of 41 amino acids at residue 116 in hVEGF. Houck *et al.*, Mol. Endocrin., **5**: 1806 (1991); Ferrara *et al.*, J. Cell. Biochem., **47**: 211 (1991); Ferrara *et al.*, Endocrine Reviews, **13**: 18 (1992); Keck *et al.*, Science, **246**: 1309 (1989); Connolly *et al.*, J. Biol. Chem., **264**: 20017 (1989); EP 370,989 published May 30, 1990.

It is now well established that angiogenesis, which involves the formation of new blood vessels from preexisting endothelium, is implicated in the pathogenesis of a variety of disorders. These include solid tumors and metastasis, atherosclerosis, retrolental fibroplasia, hemangiomas, chronic inflammation, intraocular neovascular syndromes such as proliferative retinopathies, *e.g.*, diabetic retinopathy, age-related macular degeneration (AMD), neovascular glaucoma, immune rejection of transplanted corneal tissue and other tissues, rheumatoid arthritis, and psoriasis. Folkman *et al.*, J. Biol. Chem., **267**: 10931-10934 (1992); Klagsbrun *et al.*, Annu. Rev. Physiol., **53**: 217-239 (1991); and Garner A., "Vascular diseases", In: Pathobiology of Ocular Disease. A Dynamic Approach, Garner A., Klintworth GK, eds., 2nd Edition (Marcel Dekker, NY, 1994), pp 1625-1710.

In the case of tumor growth, angiogenesis appears to be crucial for the transition from hyperplasia to neoplasia, and for providing nourishment for the growth and metastasis of the tumor. Folkman *et al.*, Nature, **339**: 58 (1989). The neovascularization allows the tumor cells to acquire a growth advantage and proliferative autonomy compared to the normal cells. A tumor usually begins as a single aberrant cell which can proliferate only to a size of a few cubic millimeters due to the distance from available capillary beds, and it can stay 'dormant' without further growth and dissemination for a long period of time. Some tumor cells then switch to the angiogenic phenotype to activate endothelial cells, which proliferate and mature into new capillary blood vessels. These newly formed blood vessels not only allow for continued growth of the primary tumor, but also for the dissemination and recolonization of metastatic tumor cells. Accordingly, a correlation has been observed between density of microvessels in tumor sections and patient survival in breast cancer as well as in several other tumors. Weidner *et al.*, N. Engl. J. Med., **324**: 1-6 (1991); Horak *et al.*, Lancet, **340**: 1120-1124 (1992); Macchiarini *et al.*, Lancet, **340**: 145-146 (1992). The precise mechanisms that control the angiogenic switch is not well understood, but it is believed that neovascularization of tumor mass results from the net balance of a multitude of angiogenesis stimulators and inhibitors (Folkman, 1995, Nat Med **1**(1):27-31).

The search for positive regulators of angiogenesis has yielded many candidates, including aFGF, bFGF, TGF- $\alpha$ , TGF- $\beta$ , HGF, TNF- $\alpha$ , angiogenin, IL-8, etc. Folkman *et al.*, J.B.C., *supra*, and Klagsbrun *et al.*, *supra*. The negative regulators so far identified include thrombospondin (Good *et al.*, Proc. Natl. Acad. Sci. USA., **87**: 6624-6628 (1990)), the 16-kilodalton N-terminal fragment of prolactin (Clapp *et al.*, Endocrinology, **133**: 1292-1299 (1993)), angiostatin (O'Reilly *et al.*, Cell, **79**: 315-328 (1994)), and endostatin. O'Reilly *et al.*, Cell, **88**: 277-285 (1996).

Work done over the last several years has established the key role of VEGF, not only in stimulating vascular endothelial cell proliferation, but also in inducing vascular permeability and angiogenesis. Ferrara *et al.*, Endocr. Rev., **18**: 4-25 (1997). The finding that the loss of even a single VEGF allele results in embryonic lethality points to an irreplaceable role played by this factor in the development and differentiation of the vascular system.

Furthermore, VEGF has been shown to be a key mediator of neovascularization associated with tumors and intraocular disorders. Ferrara *et al.*, Endocr. Rev., *supra*. The VEGF mRNA is overexpressed by the majority of human tumors examined. Berkman *et al.*, J. Clin. Invest., **91**: 153-159 (1993); Brown *et al.*, Human Pathol., **26**: 86-91 (1995); Brown *et al.*, Cancer Res., **53**: 4727-4735 (1993); Mattern *et al.*, Brit. J. Cancer, **73**: 931-934 (1996); Dvorak *et al.*, Am. J. Pathol., **146**: 1029-1039 (1995).

Also, the concentration levels of VEGF in eye fluids are highly correlated to the presence of active proliferation of blood vessels in patients with diabetic and other ischemia-related retinopathies. Aiello *et al.*, N. Engl. J. Med., **331**: 1480-1487 (1994). Furthermore, recent studies have demonstrated the localization of VEGF in choroidal neovascular membranes in patients affected by AMD. Lopez *et al.*, Invest. Ophthalmol. Vis. Sci., **37**: 855-868 (1996).

Anti-VEGF neutralizing antibodies suppress the growth of a variety of human tumor cell lines in nude mice (Kim *et al.*, Nature, **362**: 841-844 (1993); Warren *et al.*, J. Clin. Invest., **95**: 1789-1797 (1995); Borgström *et al.*, Cancer Res., **56**: 4032-4039 (1996); Melnyk *et al.*, Cancer Res., **56**: 921-924 (1996)) and also inhibit intraocular angiogenesis in models of ischemic retinal disorders. Adamis *et al.*, Arch. Ophthalmol., **114**: 66-71 (1996). Therefore, anti-VEGF monoclonal antibodies or other inhibitors of VEGF action are promising candidates for the treatment of solid tumors and various intraocular neovascular disorders. Such antibodies are described, for example, in EP 817,648 published January 14, 1998 and in WO98/45331 and WO98/45332 both published October 15, 1998.

There exist several other growth factors and mitogens, including transforming oncogenes, that are capable of rapidly inducing a complex set of genes to be expressed by certain cells. Lau and Nathans, Molecular Aspects of Cellular Regulation, **6**: 165-202 (1991). These genes, which have been named immediate-early- or early-response genes, are transcriptionally activated within minutes after contact with a growth factor or mitogen, independent of *de novo* protein synthesis. A group of these intermediate-early genes encodes secreted, extracellular proteins that are needed for coordination of complex biological processes such as differentiation and proliferation, regeneration, and wound healing. Ryseck *et al.*, Cell Growth Differ., **2**: 235-233 (1991).

Highly-related proteins that belong to this group include *cef10* (Simmons *et al.*, Proc. Natl. Acad. Sci. USA, **86**: 1178-1182 (1989)), *cyr61*, which is rapidly activated by serum- or platelet-derived growth factor (PDGF) (O'Brien *et al.*, Mol. Cell Biol., **10**: 3569-3577 (1990)), human connective tissue growth factor (CTGF) (Bradham *et al.*, J. Cell. Biol., **114**: 1285-1294 (1991)), which is secreted by human vascular endothelial cells in high levels after activation with transforming growth factor beta (TGF- $\beta$ ), exhibits PDGF-like biological and immunological activities, and competes with PDGF for a particular cell surface receptor, *fisp-12* (Ryseck *et al.*, Cell Growth Differ., **2**: 235-233 (1991)), human vascular IBP-like growth factor (VIGF) (WO 96/17931), and *nov*, normally arrested in adult kidney cells, which was found to be overexpressed in myeloblastosis-associated-virus-type-1-induced nephroblastomas. Joliet *et al.*, Mol. Cell Biol., **12**: 10-21 (1992).

The expression of these immediate-early genes acts as "third messengers" in the cascade of events triggered by growth factors. It is also thought that they are needed to integrate and coordinate complex biological processes, such as differentiation and wound healing in which cell proliferation is a common event.

As additional mitogens, insulin-like growth factor binding proteins (IGFBPs) have been shown, in complex with insulin-like growth factor (IGF), to stimulate increased binding of IGF to fibroblast and smooth muscle cell surface receptors. Clemmons *et al.*, *J. Clin. Invest.*, **77**: 1548 (1986). Inhibitory effects of IGFBP on various IGF actions *in vitro* include stimulation of glucose transport by adipocytes, sulfate incorporation by chondrocytes, and thymidine incorporation in fibroblast. Zapf *et al.*, *J. Clin. Invest.*, **63**: 1077 (1979). In addition, inhibitory effects of IGFBPs on growth factor-mediated mitogen activity in normal cells have been shown.

#### 2.4. Need for Further Treatments

In view of the role of vascular endothelial cell growth and angiogenesis in many diseases and disorders, it is desirable to have a means of reducing or inhibiting one or more of the biological effects causing these processes. It is also desirable to have a means of assaying for the presence of pathogenic polypeptides in normal and diseased conditions, and especially cancer. Further, in a specific aspect, as there is no generally applicable therapy for the treatment of cardiac hypertrophy, the identification of factors that can prevent or reduce cardiac myocyte hypertrophy is of primary importance in the development of new therapeutic strategies to inhibit pathophysiological cardiac growth. While there are several treatment modalities for various cardiovascular and oncologic disorders, there is still a need for additional therapeutic approaches.

#### 3. Summary of the Invention

The present invention provides compositions and methods for modulating (*e.g.*, promoting or inhibiting) angiogenesis and/or cardiovascularization in mammals. The present invention is based on the identification of compounds (*i.e.*, proteins) that test positive in various cardiovascular assays that test modulation (*e.g.*, promotion or inhibition) of certain biological activities. Accordingly, the compounds are believed to be useful drugs and/or drug components for the diagnosis and/or treatment (including prevention and amelioration) of disorders where such effects are desired, such as the promotion or inhibition of angiogenesis, inhibition or stimulation of vascular endothelial cell growth, stimulation of growth or proliferation of vascular endothelial cells, inhibition of tumor growth, inhibition of angiogenesis-dependent tissue growth, stimulation of angiogenesis-dependent tissue growth, inhibition of cardiac hypertrophy and stimulation of cardiac hypertrophy, *e.g.*, for the treatment of congestive heart failure. In addition, the compositions and methods of the invention provide for the diagnostic monitoring of patients undergoing clinical evaluation for the treatment of angiogenesis-related disorders, for monitoring the efficacy of compounds in clinical trials and for identifying subjects who may be predisposed to such angiogenic-related disorders.

In one embodiment, the present invention provides a composition comprising a PRO polypeptide, an agonist or antagonist thereof, or an anti-PRO antibody in admixture with a pharmaceutically acceptable carrier. In one aspect, the composition comprises a therapeutically effective amount of the polypeptide, agonist, antagonist

or antibody. In another aspect, the composition comprises a further active ingredient, namely, a cardiovascular, endothelial or angiogenic agent or an angiostatic agent, preferably an angiogenic or angiostatic agent. Preferably, the composition is sterile. The PRO polypeptide, agonist, antagonist or antibody may be administered in the form of a liquid pharmaceutical formulation, which may be preserved to achieve extended storage stability. Preserved liquid pharmaceutical formulations might contain multiple doses of PRO polypeptide, agonist, antagonist or antibody, and might, therefore, be suitable for repeated use. In a preferred embodiment, where the composition comprises an antibody, the antibody is a monoclonal antibody, an antibody fragment, a humanized antibody, or a single-chain antibody.

In a further embodiment, the present invention provides a method for preparing such a composition useful for the treatment of a cardiovascular, endothelial or angiogenic disorder comprising admixing a therapeutically effective amount of a PRO polypeptide, agonist, antagonist or antibody with a pharmaceutically acceptable carrier.

In a still further aspect, the present invention provides an article of manufacture comprising:

- (a) a composition of matter comprising a PRO polypeptide or agonist or antagonist thereof;
- (b) a container containing said composition; and

(c) a label affixed to said container, or a package insert included in said container referring to the use of said PRO polypeptide or agonist or antagonist thereof in the treatment of a cardiovascular, endothelial or angiogenic disorder, wherein the agonist or antagonist may be an antibody which binds to the PRO polypeptide. The composition may comprise a therapeutically effective amount of the PRO polypeptide or the agonist or antagonist thereof.

In another embodiment, the present invention provides a method for identifying an agonist of a PRO polypeptide comprising:

- (a) contacting cells and a test compound to be screened under conditions suitable for the induction of a cellular response normally induced by a PRO polypeptide; and
- (b) determining the induction of said cellular response to determine if the test compound is an effective agonist, wherein the induction of said cellular response is indicative of said test compound being an effective agonist.

In another embodiment, the present invention provides a method for identifying an agonist of a PRO polypeptide comprising:

- (a) contacting cells and a test compound to be screened under conditions suitable for the stimulation of cell proliferation by a PRO polypeptide; and
- (b) measuring the proliferation of said cells to determine if the test compound is an effective agonist, wherein the stimulation of cell proliferation is indicative of said test compound being an effective agonist.

In another embodiment, the invention provides a method for identifying a compound that inhibits the activity of a PRO polypeptide comprising contacting a test compound with a PRO polypeptide under conditions and for a time sufficient to allow the test compound and polypeptide to interact and determining whether the activity of the PRO polypeptide is inhibited. In a specific preferred aspect, either the test compound or the PRO polypeptide is immobilized on a solid support. In another preferred aspect, the non-immobilized component carries a detectable

label. In a preferred aspect, this method comprises the steps of:

(a) contacting cells and a test compound to be screened in the presence of a PRO polypeptide under conditions suitable for the induction of a cellular response normally induced by a PRO polypeptide; and

5 (b) determining the induction of said cellular response to determine if the test compound is an effective antagonist.

In another preferred aspect, this process comprises the steps of:

(a) contacting cells and a test compound to be screened in the presence of a PRO polypeptide under conditions suitable for the stimulation of cell proliferation by a PRO polypeptide; and

(b) measuring the proliferation of the cells to determine if the test compound is an effective antagonist.

10 In another embodiment, the invention provides a method for identifying a compound that inhibits the expression of a PRO polypeptide in cells that normally expresses the polypeptide, wherein the method comprises contacting the cells with a test compound and determining whether the expression of the PRO polypeptide is inhibited. In a preferred aspect, this method comprises the steps of:

15 (a) contacting cells and a test compound to be screened under conditions suitable for allowing expression of the PRO polypeptide; and

(b) determining the inhibition of expression of said polypeptide.

In a still further embodiment, the invention provides a compound that inhibits the expression of a PRO polypeptide, such as a compound that is identified by the methods set forth above.

20 Another aspect of the present invention is directed to an agonist or an antagonist of a PRO polypeptide which may optionally be identified by the methods described above.

One type of antagonist of a PRO polypeptide that inhibits one or more of the functions or activities of the PRO polypeptide is an antibody. Hence, in another aspect, the invention provides an isolated antibody that binds a PRO polypeptide. In a preferred aspect, the antibody is a monoclonal antibody, which preferably has non-human complementarity-determining-region (CDR) residues and human framework-region (FR) residues. The antibody  
25 may be labeled and may be immobilized on a solid support. In a further aspect, the antibody is an antibody fragment, a single-chain antibody, or a humanized antibody. Preferably, the antibody specifically binds to the polypeptide.

30 In a still further aspect, the present invention provides a method for diagnosing a disease or susceptibility to a disease which is related to a mutation in a PRO polypeptide-encoding nucleic acid sequence comprising determining the presence or absence of said mutation in the PRO polypeptide nucleic acid sequence, wherein the presence or absence of said mutation is indicative of the presence of said disease or susceptibility to said disease.

35 In a still further aspect, the invention provides a method of diagnosing a cardiovascular, endothelial or angiogenic disorder in a mammal which comprises analyzing the level of expression of a gene encoding a PRO polypeptide (a) in a test sample of tissue cells obtained from said mammal, and (b) in a control sample of known normal tissue cells of the same cell type, wherein a higher or lower expression level in the test sample as compared to the control sample is indicative of the presence of a cardiovascular, endothelial or angiogenic disorder in said mammal. The expression of a gene encoding a PRO polypeptide may optionally be accomplished by measuring

the level of mRNA or the polypeptide in the test sample as compared to the control sample.

5 In a still further aspect, the present invention provides a method of diagnosing a cardiovascular, endothelial or angiogenic disorder in a mammal which comprises detecting the presence or absence of a PRO polypeptide in a test sample of tissue cells obtained from said mammal, wherein the presence or absence of said PRO polypeptide in said test sample is indicative of the presence of a cardiovascular, endothelial or angiogenic disorder in said mammal.

10 In a still further embodiment, the invention provides a method of diagnosing a cardiovascular, endothelial or angiogenic disorder in a mammal comprising (a) contacting an anti-PRO antibody with a test sample of tissue cells obtained from the mammal, and (b) detecting the formation of a complex between the antibody and the PRO polypeptide in the test sample, wherein the formation of said complex is indicative of the presence of a cardiovascular, endothelial or angiogenic disorder in the mammal. The detection may be qualitative or quantitative, and may be performed in comparison with monitoring the complex formation in a control sample of known normal tissue cells of the same cell type. A larger or smaller quantity of complexes formed in the test sample indicates the presence of a cardiovascular, endothelial or angiogenic dysfunction in the mammal from which the test tissue cells were obtained. The antibody preferably carries a detectable label. Complex formation can be monitored, for example, by light microscopy, flow cytometry, fluorimetry, or other techniques known in the art. The test sample is usually obtained from an individual suspected to have a cardiovascular, endothelial or angiogenic disorder.

15 In another embodiment, the invention provides a method for determining the presence of a PRO polypeptide in a sample comprising exposing a sample suspected of containing the PRO polypeptide to an anti-PRO antibody and determining binding of said antibody to a component of said sample. In a specific aspect, the sample comprises a cell suspected of containing the PRO polypeptide and the antibody binds to the cell. The antibody is preferably detectably labeled and/or bound to a solid support.

20 In further aspects, the invention provides a cardiovascular, endothelial or angiogenic disorder diagnostic kit comprising an anti-PRO antibody and a carrier in suitable packaging. Preferably, such kit further comprises instructions for using said antibody to detect the presence of the PRO polypeptide. Preferably, the carrier is a buffer, for example. Preferably, the cardiovascular, endothelial or angiogenic disorder is cancer.

25 In yet another embodiment, the present invention provides a method for treating a cardiovascular, endothelial or angiogenic disorder in a mammal comprising administering to the mammal an effective amount of a PRO polypeptide. Preferably, the disorder is cardiac hypertrophy, trauma such as wounds or burns, or a type of cancer. In a further aspect, the mammal is further exposed to angioplasty or a drug that treats cardiovascular, endothelial or angiogenic disorders such as ACE inhibitors or chemotherapeutic agents if the cardiovascular, endothelial or angiogenic disorder is a type of cancer. Preferably, the mammal is human, preferably one who is at risk of developing cardiac hypertrophy and more preferably has suffered myocardial infarction.

30 In another preferred aspect, the cardiac hypertrophy is characterized by the presence of an elevated level of  $\text{PGF}_{2\alpha}$ . Alternatively, the cardiac hypertrophy may be induced by myocardial infarction, wherein preferably the administration of the PRO polypeptide is initiated within 48 hours, more preferably within 24 hours, following myocardial infarction.



In another preferred embodiment, the cardiovascular, endothelial or angiogenic disorder is cardiac hypertrophy and said PRO polypeptide is administered together with a cardiovascular, endothelial or angiogenic agent. The preferred cardiovascular, endothelial or angiogenic agent for this purpose is selected from the group consisting of an antihypertensive drug, an ACE inhibitor, an endothelin receptor antagonist and a thrombolytic agent. If a thrombolytic agent is administered, preferably the PRO polypeptide is administered following administration of such agent. More preferably, the thrombolytic agent is recombinant human tissue plasminogen activator.

In another preferred aspect, the cardiovascular, endothelial or angiogenic disorder is cardiac hypertrophy and the PRO polypeptide is administered following primary angioplasty for the treatment of acute myocardial infarction, preferably wherein the mammal is further exposed to angioplasty or a cardiovascular, endothelial, or angiogenic agent.

In another preferred embodiment, the cardiovascular, endothelial or angiogenic disorder is a cancer and the PRO polypeptide is administered in combination with a chemotherapeutic agent, a growth inhibitory agent or a cytotoxic agent.

In a further embodiment, the invention provides a method for treating a cardiovascular, endothelial or angiogenic disorder in a mammal comprising administering to the mammal an effective amount of a PRO polypeptide agonist, antagonist or anti-PRO antibody. Preferably, the cardiovascular, endothelial or angiogenic disorder is cardiac hypertrophy, trauma, a cancer, or age-related macular degeneration. Also preferred is where the mammal is human, and where an effective amount of an angiogenic or angiostatic agent is administered in conjunction with the agonist, antagonist or anti-PRO antibody.

In still further embodiments, the invention provides a method for treating a cardiovascular, endothelial or angiogenic disorder in a mammal that suffers therefrom comprising administering to the mammal a nucleic acid molecule that codes for either (a) a PRO polypeptide, (b) an agonist of a PRO polypeptide or (c) an antagonist of a PRO polypeptide, wherein said agonist or antagonist may be an anti-PRO antibody. In a preferred embodiment, the mammal is human. In another preferred embodiment, the gene is administered via *ex vivo* gene therapy. In a further preferred embodiment, the gene is comprised within a vector, more preferably an adenoviral, adeno-associated viral, lentiviral, or retroviral vector.

In yet another aspect, the invention provides a recombinant retroviral particle comprising a retroviral vector consisting essentially of a promoter, nucleic acid encoding (a) a PRO polypeptide, (b) an agonist polypeptide of a PRO polypeptide, or (c) an antagonist polypeptide of a PRO polypeptide, and a signal sequence for cellular secretion of the polypeptide, wherein the retroviral vector is in association with retroviral structural proteins. Preferably, the signal sequence is from a mammal, such as from a native PRO polypeptide.

In a still further embodiment, the invention supplies an *ex vivo* producer cell comprising a nucleic acid construct that expresses retroviral structural proteins and also comprises a retroviral vector consisting essentially of a promoter, nucleic acid encoding (a) a PRO polypeptide, (b) an agonist polypeptide of a PRO polypeptide or (c) an antagonist polypeptide of a PRO polypeptide, and a signal sequence for cellular secretion of the polypeptide, wherein said producer cell packages the retroviral vector in association with the structural proteins to produce

recombinant retroviral particles.

In yet another embodiment, the invention provides a method for inhibiting endothelial cell growth in a mammal comprising administering to the mammal (a) a PRO polypeptide, (b) an agonist of a PRO polypeptide, or (c) an antagonist of a PRO polypeptide, wherein endothelial cell growth in said mammal is inhibited, and wherein said agonist or antagonist may be an anti-PRO antibody. Preferably, the mammal is human and the endothelial cell growth is associated with a tumor or a retinal disorder.

In yet another embodiment, the invention provides a method for stimulating endothelial cell growth in a mammal comprising administering to the mammal (a) a PRO polypeptide, (b) an agonist of a PRO polypeptide, or (c) an antagonist of a PRO polypeptide, wherein endothelial cell growth in said mammal is stimulated, and wherein said agonist or antagonist may be an anti-PRO antibody. Preferably, the mammal is human.

In yet another embodiment, the invention provides a method for inhibiting cardiac hypertrophy in a mammal comprising administering to the mammal (a) a PRO polypeptide, (b) an agonist of a PRO polypeptide, or (c) an antagonist of a PRO polypeptide, wherein cardiac hypertrophy in said mammal is inhibited, and wherein said agonist or antagonist may be an anti-PRO antibody. Preferably, the mammal is human and the cardiac hypertrophy has been induced by myocardial infarction.

In yet another embodiment, the invention provides a method for stimulating cardiac hypertrophy in a mammal comprising administering to the mammal (a) a PRO polypeptide, (b) an agonist of a PRO polypeptide, or (c) an antagonist of a PRO polypeptide, wherein cardiac hypertrophy in said mammal is stimulated, and wherein said agonist or antagonist may be an anti-PRO antibody. Preferably, the mammal is human who suffers from congestive heart failure.

In yet another embodiment, the invention provides a method for inhibiting angiogenesis induced by a PRO polypeptide in a mammal comprising administering a therapeutically effective amount of an anti-PRO antibody to the mammal. Preferably, the mammal is a human, and more preferably the mammal has a tumor or a retinal disorder.

In yet another embodiment, the invention provides a method for stimulating angiogenesis induced by a PRO polypeptide in a mammal comprising administering a therapeutically effective amount of a PRO polypeptide to the mammal. Preferably, the mammal is a human, and more preferably angiogenesis would promote tissue regeneration or wound healing.

In yet another embodiment, the invention provides a method for modulating (e.g., inhibiting or stimulating) endothelial cell growth in a mammal comprising administering to the mammal a PRO21, PRO181, PRO205, PRO214, PRO221, PRO229, PRO231, PRO238, PRO241, PRO247, PRO256, PRO258, PRO263, PRO265, PRO295, PRO321, PRO322, PRO337, PRO363, PRO365, PRO444, PRO533, PRO697, PRO720, PRO725, PRO771, PRO788, PRO791, PRO819, PRO827, PRO828, PRO836, PRO846, PRO865, PRO1005, PRO1006, PRO1007, PRO1025, PRO1029, PRO1054, PRO1071, PRO1075, PRO1079, PRO1080, PRO1114, PRO1131, PRO1155, PRO1160, PRO1184, PRO1186, PRO1190, PRO1192, PRO1195, PRO1244, PRO1272, PRO1273, PRO1274, PRO1279, PRO1283, PRO1286, PRO1306, PRO1309, PRO1325, PRO1329, PRO1347, PRO1356, PRO1376, PRO1382, PRO1411, PRO1412, PRO1419, PRO1474, PRO1477, PRO1488, PRO1508, PRO1550,

PRO1556, PRO1760, PRO1782, PRO1787, PRO1801, PRO1868, PRO1887, PRO1890, PRO3438, PRO3444, PRO4302, PRO4324, PRO4333, PRO4341, PRO4342, PRO4353, PRO4354, PRO4356, PRO4371, PRO4405, PRO4408, PRO4422, PRO4425, PRO4499, PRO5723, PRO5725, PRO5737, PRO5776, PRO6006, PRO6029, PRO6071, PRO7436, PRO9771, PRO9821, PRO9873, PRO10008, PRO10096, PRO19670, PRO20040, PRO20044, PRO21055, PRO21384 or PRO28631 polypeptide, agonist or antagonist thereof, wherein endothelial cell growth in said mammal is modulated.

In yet another embodiment, the invention provides a method for modulating (*e.g.*, inhibiting or stimulating) smooth muscle cell growth in a mammal comprising administering to the mammal a PRO162, PRO181, PRO182, PRO195, PRO204, PRO221, PRO230, PRO256, PRO258, PRO533, PRO697, PRO725, PRO738, PRO826, PRO836, PRO840, PRO846, PRO865, PRO982, PRO1025, PRO1029, PRO1071, PRO1080, PRO1083, PRO1134, PRO1160, PRO1182, PRO1184, PRO1186, PRO1192, PRO1265, PRO1274, PRO1279, PRO1283, PRO1306, PRO1308, PRO1309, PRO1325, PRO1337, PRO1338, PRO1343, PRO1376, PRO1387, PRO1411, PRO1412, PRO1415, PRO1434, PRO1474, PRO1488, PRO1550, PRO1556, PRO1567, PRO1600, PRO1754, PRO1758, PRO1760, PRO1787, PRO1865, PRO1868, PRO1917, PRO1928, PRO3438, PRO3562, PRO4302, PRO4333, PRO4345, PRO4353, PRO4354, PRO4405, PRO4408, PRO4430, PRO4503, PRO5725, PRO6714, PRO9771, PRO9820, PRO9940, PRO10096, PRO21055, PRO21184 or PRO21366 polypeptide, agonist or antagonist thereof, wherein endothelial cell growth in said mammal is modulated.

In yet another embodiment, the invention provides a method for modulating (*e.g.*, inducing or reducing) cardiac hypertrophy in a mammal comprising administering to the mammal a PRO21 polypeptide, agonist or antagonist thereof, wherein cardiac hypertrophy in said mammal is modulated.

In yet another embodiment, the invention provides a method for modulating (*e.g.*, inducing or reducing) endothelial cell apoptosis in a mammal comprising administering to the mammal a PRO4302 polypeptide, agonist or antagonist thereof, wherein cardiac hypertrophy in said mammal is modulated.

In yet another embodiment, the invention provides a method for modulating (*e.g.*, stimulating or inhibiting) angiogenesis in a mammal comprising administering a therapeutically effective amount of a PRO1376 or PRO1449 polypeptide, agonist or antagonist thereof to the mammal, wherein said angiogenesis is modulated.

In yet another embodiment, the invention provides a method for modulating (*e.g.*, inducing or reducing) angiogenesis by modulating (*e.g.*, inducing or reducing) endothelial cell tube formation in a mammal comprising administering to the mammal a PRO178, PRO195, PRO228, PRO301, PRO302, PRO532, PRO724, PRO730, PRO734, PRO793, PRO871, PRO938, PRO1012, PRO1120, PRO1139, PRO1198, PRO1287, PRO1361, PRO1864, PRO1873, PRO2010, PRO3579, PRO4313, PRO4527, PRO4538, PRO4553, PRO4995, PRO5730, PRO6008, PRO7223, PRO7248 or PRO7261 polypeptide, agonist or antagonist thereof, wherein endothelial cell tube formation in said mammal is modulated.

In other embodiments of the present invention, the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence that encodes a PRO polypeptide.

In one aspect, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97% or 98%

nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule encoding a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In other aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97% or 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule comprising the coding sequence of a full-length PRO polypeptide cDNA as disclosed herein, the coding sequence of a PRO polypeptide lacking the signal peptide as disclosed herein, the coding sequence of an extracellular domain of a transmembrane PRO polypeptide, with or without the signal peptide, as disclosed herein or the coding sequence of any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In a further aspect, the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence having at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97% or 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule that encodes the same mature polypeptide encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein, or (b) the complement of the DNA molecule of (a).

Another aspect of the present invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated, or is complementary to such encoding nucleotide sequence, wherein the transmembrane domain(s) of such polypeptide are disclosed herein. Therefore, soluble extracellular domains of the herein described PRO polypeptides are contemplated.

Another embodiment is directed to fragments of a PRO polypeptide coding sequence, or the complement thereof, that may find use as, for example, hybridization probes, for encoding fragments of a PRO polypeptide that may optionally encode a polypeptide comprising a binding site for an anti-PRO antibody or as antisense oligonucleotide probes. Such nucleic acid fragments are usually at least about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 350, 400, 450, 500, 600, 700 or 800 nucleotides in length and alternatively at least about 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length. It is noted that novel fragments of a PRO polypeptide-encoding nucleotide sequence may be determined in a routine manner by aligning the PRO polypeptide-encoding nucleotide sequence with other known nucleotide sequences using any of a number of well known sequence alignment programs and determining which PRO polypeptide-encoding nucleotide sequence fragment(s) are novel. All of such PRO polypeptide-encoding nucleotide sequences are contemplated herein. Also contemplated are the PRO polypeptide fragments encoded by these nucleotide molecule fragments, preferably those PRO polypeptide fragments that comprise a binding site for an anti-PRO antibody.

In another embodiment, the invention provides an isolated PRO polypeptide encoded by any of the isolated

nucleic acid sequences hereinabove identified.

In a certain aspect, the invention provides an isolated PRO polypeptide comprising an amino acid sequence having at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97% or 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of the full-length amino acid sequence as disclosed herein.

In a further aspect, the invention provides an isolated PRO polypeptide comprising an amino acid sequence having at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97% or 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to an amino acid sequence encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein.

In a specific aspect, the invention provides an isolated PRO polypeptide without the N-terminal signal sequence and/or the initiating methionine and that is encoded by a nucleotide sequence that encodes such an amino acid sequence as hereinbefore described. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

Another aspect of the invention provides an isolated PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

In yet another embodiment, the invention provides agonists and antagonists of a native PRO polypeptide as defined herein. In a particular embodiment, the agonist or antagonist is an anti-PRO antibody or a small molecule.

In a further embodiment, the invention provides a method of identifying agonists or antagonists to a PRO polypeptide which comprise contacting the PRO polypeptide with a candidate molecule and monitoring a biological activity mediated by said PRO polypeptide. Preferably, the PRO polypeptide is a native PRO polypeptide.

In a still further embodiment, the invention provides a composition of matter comprising a PRO polypeptide, or an agonist or antagonist of a PRO polypeptide as herein described, or an anti-PRO antibody, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier.

Another embodiment of the present invention is directed to the use of a PRO polypeptide, or an agonist or antagonist thereof as hereinbefore described, or an anti-PRO antibody, for the preparation of a medicament useful in the treatment of a condition which is responsive to the PRO polypeptide, an agonist or antagonist thereof or an anti-PRO antibody.

In additional embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described polypeptides. Host cells comprising any such vector are also provided. By way of example, the host cells may be CHO cells, *E. coli*, yeast, or Baculovirus-infected insect cells. A process for producing any of the herein described polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired polypeptide and recovering the desired polypeptide from the cell culture.

In other embodiments, the invention provides chimeric molecules comprising any of the herein described polypeptides fused to a heterologous polypeptide or amino acid sequence. Example of such chimeric molecules comprise any of the herein described polypeptides fused to an epitope tag sequence or a Fc region of an immunoglobulin.

In yet another embodiment, the invention provides an antibody which specifically binds to any of the above or below described polypeptides. Optionally, the antibody is a monoclonal antibody, humanized antibody, antibody fragment or single-chain antibody.

In yet other embodiments, the invention provides oligonucleotide probes useful for isolating genomic and cDNA nucleotide sequences or as antisense probes, wherein those probes may be derived from any of the above or below described nucleotide sequences.

#### 4. Brief Description of the Drawings

Figure 1 shows a nucleotide sequence (SEQ ID NO:1) of a native sequence PRO181 cDNA, wherein SEQ ID NO:1 is a clone designated herein as "DNA23330-1390".

Figure 2 shows the amino acid sequence (SEQ ID NO:2) derived from the coding sequence of SEQ ID NO:1 shown in Figure 1.

Figure 3 shows a nucleotide sequence (SEQ ID NO:3) of a native sequence PRO178 cDNA, wherein SEQ ID NO:3 is a clone designated herein as "DNA23339-1130".

Figure 4 shows the amino acid sequence (SEQ ID NO:4) derived from the coding sequence of SEQ ID NO:3 shown in Figure 3.

Figure 5 shows a nucleotide sequence (SEQ ID NO:5) of a native sequence PRO444 cDNA, wherein SEQ ID NO:5 is a clone designated herein as "DNA26846-1397".

Figure 6 shows the amino acid sequence (SEQ ID NO:6) derived from the coding sequence of SEQ ID NO:5 shown in Figure 5.

Figure 7 shows a nucleotide sequence (SEQ ID NO:7) of a native sequence PRO195 cDNA, wherein SEQ ID NO:7 is a clone designated herein as "DNA26847-1395".

Figure 8 shows the amino acid sequence (SEQ ID NO:8) derived from the coding sequence of SEQ ID NO:7 shown in Figure 7.

Figure 9 shows a nucleotide sequence (SEQ ID NO:9) of a native sequence PRO182 cDNA, wherein SEQ ID NO:9 is a clone designated herein as "DNA27865-1091".

Figure 10 shows the amino acid sequence (SEQ ID NO:10) derived from the coding sequence of SEQ ID

NO:9 shown in Figure 9.

Figure 11 shows a nucleotide sequence (SEQ ID NO:11) of a native sequence PRO205 cDNA, wherein SEQ ID NO:11 is a clone designated herein as "DNA30868-1156".

5 Figure 12 shows the amino acid sequence (SEQ ID NO:12) derived from the coding sequence of SEQ ID NO:11 shown in Figure 11.

Figure 13 shows a nucleotide sequence (SEQ ID NO:13) of a native sequence PRO204 cDNA, wherein SEQ ID NO:13 is a clone designated herein as "DNA30871-1157".

Figure 14 shows the amino acid sequence (SEQ ID NO:14) derived from the coding sequence of SEQ ID NO:13 shown in Figure 13.

10 Figure 15 shows a nucleotide sequence (SEQ ID NO:15) of a native sequence PRO1873 cDNA, wherein SEQ ID NO:15 is a clone designated herein as "DNA30880".

Figure 16 shows the amino acid sequence (SEQ ID NO:16) derived from the coding sequence of SEQ ID NO:15 shown in Figure 15.

15 Figure 17 shows a nucleotide sequence (SEQ ID NO:17) of a native sequence PRO214 cDNA, wherein SEQ ID NO:17 is a clone designated herein as "DNA32286-1191".

Figure 18 shows the amino acid sequence (SEQ ID NO:18) derived from the coding sequence of SEQ ID NO:17 shown in Figure 17.

Figure 19 shows a nucleotide sequence (SEQ ID NO:19) of a native sequence PRO221 cDNA, wherein SEQ ID NO:19 is a clone designated herein as "DNA33089-1132".

20 Figure 20 shows the amino acid sequence (SEQ ID NO:20) derived from the coding sequence of SEQ ID NO:19 shown in Figure 19.

Figure 21 shows a nucleotide sequence (SEQ ID NO:21) of a native sequence PRO228 cDNA, wherein SEQ ID NO:21 is a clone designated herein as "DNA33092-1202".

25 Figure 22 shows the amino acid sequence (SEQ ID NO:22) derived from the coding sequence of SEQ ID NO:21 shown in Figure 21.

Figure 23 shows a nucleotide sequence (SEQ ID NO:23) of a native sequence PRO229 cDNA, wherein SEQ ID NO:23 is a clone designated herein as "DNA33100-1159".

Figure 24 shows the amino acid sequence (SEQ ID NO:24) derived from the coding sequence of SEQ ID NO:23 shown in Figure 23.

30 Figure 25 shows a nucleotide sequence (SEQ ID NO:25) of a native sequence PRO230 cDNA, wherein SEQ ID NO:25 is a clone designated herein as "DNA33223-1136".

Figure 26 shows the amino acid sequence (SEQ ID NO:26) derived from the coding sequence of SEQ ID NO:25 shown in Figure 25.

35 Figure 27 shows a nucleotide sequence (SEQ ID NO:27) of a native sequence PRO7223 cDNA, wherein SEQ ID NO:27 is a clone designated herein as "DNA34385".

Figure 28 shows the amino acid sequence (SEQ ID NO:28) derived from the coding sequence of SEQ ID NO:27 shown in Figure 27.

Figure 29 shows a nucleotide sequence (SEQ ID NO:29) of a native sequence PRO241 cDNA, wherein SEQ ID NO:29 is a clone designated herein as "DNA34392-1170".

Figure 30 shows the amino acid sequence (SEQ ID NO:30) derived from the coding sequence of SEQ ID NO:29 shown in Figure 29.

5        Figure 31 shows a nucleotide sequence (SEQ ID NO:31) of a native sequence PRO263 cDNA, wherein SEQ ID NO:31 is a clone designated herein as "DNA34431-1177".

Figure 32 shows the amino acid sequence (SEQ ID NO:32) derived from the coding sequence of SEQ ID NO:31 shown in Figure 31.

10       Figure 33 shows a nucleotide sequence (SEQ ID NO:33) of a native sequence PRO321 cDNA, wherein SEQ ID NO:33 is a clone designated herein as "DNA34433-1308".

Figure 34 shows the amino acid sequence (SEQ ID NO:34) derived from the coding sequence of SEQ ID NO:33 shown in Figure 33.

Figure 35 shows a nucleotide sequence (SEQ ID NO:35) of a native sequence PRO231 cDNA, wherein SEQ ID NO:35 is a clone designated herein as "DNA34434-1139".

15       Figure 36 shows the amino acid sequence (SEQ ID NO:36) derived from the coding sequence of SEQ ID NO:35 shown in Figure 35.

Figure 37 shows a nucleotide sequence (SEQ ID NO:37) of a native sequence PRO238 cDNA, wherein SEQ ID NO:37 is a clone designated herein as "DNA35600-1162".

20       Figure 38 shows the amino acid sequence (SEQ ID NO:38) derived from the coding sequence of SEQ ID NO:37 shown in Figure 37.

Figure 39 shows a nucleotide sequence (SEQ ID NO:39) of a native sequence PRO247 cDNA, wherein SEQ ID NO:39 is a clone designated herein as "DNA35673-1201".

Figure 40 shows the amino acid sequence (SEQ ID NO:40) derived from the coding sequence of SEQ ID NO:39 shown in Figure 39.

25       Figure 41 shows a nucleotide sequence (SEQ ID NO:41) of a native sequence PRO256 cDNA, wherein SEQ ID NO:41 is a clone designated herein as "DNA35880-1160".

Figure 42 shows the amino acid sequence (SEQ ID NO:42) derived from the coding sequence of SEQ ID NO:41 shown in Figure 41.

30       Figure 43 shows a nucleotide sequence (SEQ ID NO:43) of a native sequence PRO258 cDNA, wherein SEQ ID NO:43 is a clone designated herein as "DNA35918-1174".

Figure 44 shows the amino acid sequence (SEQ ID NO:44) derived from the coding sequence of SEQ ID NO:43 shown in Figure 43.

Figure 45 shows a nucleotide sequence (SEQ ID NO:45) of a native sequence PRO265 cDNA, wherein SEQ ID NO:45 is a clone designated herein as "DNA36350-1158".

35       Figure 46 shows the amino acid sequence (SEQ ID NO:46) derived from the coding sequence of SEQ ID NO:45 shown in Figure 45.

Figure 47 shows a nucleotide sequence (SEQ ID NO:47) of a native sequence PRO21 cDNA, wherein SEQ



ID NO:47 is a clone designated herein as "DNA36638-1056".

Figure 48 shows the amino acid sequence (SEQ ID NO:48) derived from the coding sequence of SEQ ID NO:47 shown in Figure 47.

5 Figure 49 shows a nucleotide sequence (SEQ ID NO:49) of a native sequence PRO295 cDNA, wherein SEQ ID NO:49 is a clone designated herein as "DNA38268-1188".

Figure 50 shows the amino acid sequence (SEQ ID NO:50) derived from the coding sequence of SEQ ID NO:49 shown in Figure 49.

Figure 51 shows a nucleotide sequence (SEQ ID NO:51) of a native sequence PRO302 cDNA, wherein SEQ ID NO:51 is a clone designated herein as "DNA40370-1217".

10 Figure 52 shows the amino acid sequence (SEQ ID NO:52) derived from the coding sequence of SEQ ID NO:51 shown in Figure 51.

Figure 53 shows a nucleotide sequence (SEQ ID NO:53) of a native sequence PRO301 cDNA, wherein SEQ ID NO:53 is a clone designated herein as "DNA40628-1216".

15 Figure 54 shows the amino acid sequence (SEQ ID NO:54) derived from the coding sequence of SEQ ID NO:53 shown in Figure 53.

Figure 55 shows a nucleotide sequence (SEQ ID NO:55) of a native sequence PRO337 cDNA, wherein SEQ ID NO:55 is a clone designated herein as "DNA43316-1237".

Figure 56 shows the amino acid sequence (SEQ ID NO:56) derived from the coding sequence of SEQ ID NO:55 shown in Figure 55.

20 Figure 57 shows a nucleotide sequence (SEQ ID NO:57) of a native sequence PRO7248 cDNA, wherein SEQ ID NO:57 is a clone designated herein as "DNA44195".

Figure 58 shows the amino acid sequence (SEQ ID NO:58) derived from the coding sequence of SEQ ID NO:57 shown in Figure 57.

25 Figure 59 shows a nucleotide sequence (SEQ ID NO:59) of a native sequence PRO846 cDNA, wherein SEQ ID NO:59 is a clone designated herein as "DNA44196-1353".

Figure 60 shows the amino acid sequence (SEQ ID NO:60) derived from the coding sequence of SEQ ID NO:59 shown in Figure 59.

Figure 61 shows a nucleotide sequence (SEQ ID NO:61) of a native sequence PRO1864 cDNA, wherein SEQ ID NO:61 is a clone designated herein as "DNA45409-2511".

30 Figure 62 shows the amino acid sequence (SEQ ID NO:62) derived from the coding sequence of SEQ ID NO:61 shown in Figure 61.

Figure 63 shows a nucleotide sequence (SEQ ID NO:63) of a native sequence PRO363 cDNA, wherein SEQ ID NO:63 is a clone designated herein as "DNA45419-1252".

35 Figure 64 shows the amino acid sequence (SEQ ID NO:64) derived from the coding sequence of SEQ ID NO:63 shown in Figure 63.

Figure 65 shows a nucleotide sequence (SEQ ID NO:65) of a native sequence PRO730 cDNA, wherein SEQ ID NO:65 is a clone designated herein as "DNA45624-1400".

Figure 66 shows the amino acid sequence (SEQ ID NO:66) derived from the coding sequence of SEQ ID NO:65 shown in Figure 65.

Figure 67 shows a nucleotide sequence (SEQ ID NO:67) of a native sequence PRO365 cDNA, wherein SEQ ID NO:67 is a clone designated herein as "DNA46777-1253".

5        Figure 68 shows the amino acid sequence (SEQ ID NO:68) derived from the coding sequence of SEQ ID NO:67 shown in Figure 67.

Figure 69 shows a nucleotide sequence (SEQ ID NO:69) of a native sequence PRO532 cDNA, wherein SEQ ID NO:69 is a clone designated herein as "DNA48335".

10       Figure 70 shows the amino acid sequence (SEQ ID NO:70) derived from the coding sequence of SEQ ID NO:69 shown in Figure 69.

Figure 71 shows a nucleotide sequence (SEQ ID NO:71) of a native sequence PRO322 cDNA, wherein SEQ ID NO:71 is a clone designated herein as "DNA48336-1309".

Figure 72 shows the amino acid sequence (SEQ ID NO:72) derived from the coding sequence of SEQ ID NO:71 shown in Figure 71.

15       Figure 73 shows a nucleotide sequence (SEQ ID NO:73) of a native sequence PRO1120 cDNA, wherein SEQ ID NO:73 is a clone designated herein as "DNA48606-1479".

Figure 74 shows the amino acid sequence (SEQ ID NO:74) derived from the coding sequence of SEQ ID NO:73 shown in Figure 73.

20       Figure 75 shows a nucleotide sequence (SEQ ID NO:75) of a native sequence PRO7261 cDNA, wherein SEQ ID NO:75 is a clone designated herein as "DNA49149".

Figure 76 shows the amino acid sequence (SEQ ID NO:76) derived from the coding sequence of SEQ ID NO:75 shown in Figure 75.

Figure 77 shows a nucleotide sequence (SEQ ID NO:77) of a native sequence PRO533 cDNA, wherein SEQ ID NO:77 is a clone designated herein as "DNA49435-1219".

25       Figure 78 shows the amino acid sequence (SEQ ID NO:78) derived from the coding sequence of SEQ ID NO:77 shown in Figure 77.

Figure 79 shows a nucleotide sequence (SEQ ID NO:79) of a native sequence PRO724 cDNA, wherein SEQ ID NO:79 is a clone designated herein as "DNA49631-1328".

30       Figure 80 shows the amino acid sequence (SEQ ID NO:80) derived from the coding sequence of SEQ ID NO:79 shown in Figure 79.

Figure 81 shows a nucleotide sequence (SEQ ID NO:81) of a native sequence PRO734 cDNA, wherein SEQ ID NO:81 is a clone designated herein as "DNA49817".

Figure 82 shows the amino acid sequence (SEQ ID NO:82) derived from the coding sequence of SEQ ID NO:81 shown in Figure 81.

35       Figure 83 shows a nucleotide sequence (SEQ ID NO:83) of a native sequence PRO771 cDNA, wherein SEQ ID NO:83 is a clone designated herein as "DNA49829-1346".

Figure 84 shows the amino acid sequence (SEQ ID NO:84) derived from the coding sequence of SEQ ID

NO:83 shown in Figure 83.

Figure 85 shows a nucleotide sequence (SEQ ID NO:85) of a native sequence PRO2010 cDNA, wherein SEQ ID NO:85 is a clone designated herein as "DNA50792".

5 Figure 86 shows the amino acid sequence (SEQ ID NO:86) derived from the coding sequence of SEQ ID NO:85 shown in Figure 85.

Figure 87 shows a nucleotide sequence (SEQ ID NO:87) of a native sequence PRO871 cDNA, wherein SEQ ID NO:87 is a clone designated herein as "DNA50919-1361".

Figure 88 shows the amino acid sequence (SEQ ID NO:88) derived from the coding sequence of SEQ ID NO:87 shown in Figure 87.

10 Figure 89 shows a nucleotide sequence (SEQ ID NO:89) of a native sequence PRO697 cDNA, wherein SEQ ID NO:89 is a clone designated herein as "DNA50920-1325".

Figure 90 shows the amino acid sequence (SEQ ID NO:90) derived from the coding sequence of SEQ ID NO:89 shown in Figure 89.

15 Figure 91 shows a nucleotide sequence (SEQ ID NO:91) of a native sequence PRO1083 cDNA, wherein SEQ ID NO:91 is a clone designated herein as "DNA50921-1458".

Figure 92 shows the amino acid sequence (SEQ ID NO:92) derived from the coding sequence of SEQ ID NO:91 shown in Figure 91.

Figure 93 shows a nucleotide sequence (SEQ ID NO:93) of a native sequence PRO725 cDNA, wherein SEQ ID NO:93 is a clone designated herein as "DNA52758-1399".

20 Figure 94 shows the amino acid sequence (SEQ ID NO:94) derived from the coding sequence of SEQ ID NO:93 shown in Figure 93.

Figure 95 shows a nucleotide sequence (SEQ ID NO:95) of a native sequence PRO720 cDNA, wherein SEQ ID NO:95 is a clone designated herein as "DNA53517-1366-1".

25 Figure 96 shows the amino acid sequence (SEQ ID NO:96) derived from the coding sequence of SEQ ID NO:95 shown in Figure 95.

Figure 97 shows a nucleotide sequence (SEQ ID NO:97) of a native sequence PRO738 cDNA, wherein SEQ ID NO:97 is a clone designated herein as "DNA53915-1258".

Figure 98 shows the amino acid sequence (SEQ ID NO:98) derived from the coding sequence of SEQ ID NO:97 shown in Figure 97.

30 Figure 99 shows a nucleotide sequence (SEQ ID NO:99) of a native sequence PRO865 cDNA, wherein SEQ ID NO:99 is a clone designated herein as "DNA53974-1401".

Figure 100 shows the amino acid sequence (SEQ ID NO:100) derived from the coding sequence of SEQ ID NO:99 shown in Figure 99.

35 Figure 101 shows a nucleotide sequence (SEQ ID NO:101) of a native sequence PRO840 cDNA, wherein SEQ ID NO:101 is a clone designated herein as "DNA53987-1438".

Figure 102 shows the amino acid sequence (SEQ ID NO:102) derived from the coding sequence of SEQ ID NO:101 shown in Figure 101.

Figure 103 shows a nucleotide sequence (SEQ ID NO:103) of a native sequence PRO1080 cDNA, wherein SEQ ID NO:103 is a clone designated herein as "DNA56047-1456".

Figure 104 shows the amino acid sequence (SEQ ID NO:104) derived from the coding sequence of SEQ ID NO:103 shown in Figure 103.

5        Figure 105 shows a nucleotide sequence (SEQ ID NO:105) of a native sequence PRO1079 cDNA, wherein SEQ ID NO:105 is a clone designated herein as "DNA56050-1455".

Figure 106 shows the amino acid sequence (SEQ ID NO:106) derived from the coding sequence of SEQ ID NO:105 shown in Figure 105.

10        Figure 107 shows a nucleotide sequence (SEQ ID NO:107) of a native sequence PRO793 cDNA, wherein SEQ ID NO:107 is a clone designated herein as "DNA56110-1437".

Figure 108 shows the amino acid sequence (SEQ ID NO:108) derived from the coding sequence of SEQ ID NO:107 shown in Figure 107.

Figure 109 shows a nucleotide sequence (SEQ ID NO:109) of a native sequence PRO788 cDNA, wherein SEQ ID NO:109 is a clone designated herein as "DNA56405-1357".

15        Figure 110 shows the amino acid sequence (SEQ ID NO:110) derived from the coding sequence of SEQ ID NO:109 shown in Figure 109.

Figure 111 shows a nucleotide sequence (SEQ ID NO:111) of a native sequence PRO938 cDNA, wherein SEQ ID NO:111 is a clone designated herein as "DNA56433-1406".

20        Figure 112 shows the amino acid sequence (SEQ ID NO:112) derived from the coding sequence of SEQ ID NO:111 shown in Figure 111.

Figure 113 shows a nucleotide sequence (SEQ ID NO:113) of a native sequence PRO1012 cDNA, wherein SEQ ID NO:113 is a clone designated herein as "DNA56439-1376".

Figure 114 shows the amino acid sequence (SEQ ID NO:114) derived from the coding sequence of SEQ ID NO:113 shown in Figure 113.

25        Figure 115 shows a nucleotide sequence (SEQ ID NO:115) of a native sequence PRO1477 cDNA, wherein SEQ ID NO:115 is a clone designated herein as "DNA56529-1647".

Figure 116 shows the amino acid sequence (SEQ ID NO:116) derived from the coding sequence of SEQ ID NO:115 shown in Figure 115.

30        Figure 117 shows a nucleotide sequence (SEQ ID NO:117) of a native sequence PRO1134 cDNA, wherein SEQ ID NO:117 is a clone designated herein as "DNA56865-1491".

Figure 118 shows the amino acid sequence (SEQ ID NO:118) derived from the coding sequence of SEQ ID NO:117 shown in Figure 117.

Figure 119 shows a nucleotide sequence (SEQ ID NO:119) of a native sequence PRO162 cDNA, wherein SEQ ID NO:119 is a clone designated herein as "DNA56965-1356".

35        Figure 120 shows the amino acid sequence (SEQ ID NO:120) derived from the coding sequence of SEQ ID NO:119 shown in Figure 119.

Figure 121 shows a nucleotide sequence (SEQ ID NO:121) of a native sequence PRO1114 cDNA, wherein

SEQ ID NO:121 is a clone designated herein as "DNA57033-1403-1".

Figure 122 shows the amino acid sequence (SEQ ID NO:122) derived from the coding sequence of SEQ ID NO:121 shown in Figure 121.

5       Figure 123 shows a nucleotide sequence (SEQ ID NO:123) of a native sequence PRO828 cDNA, wherein SEQ ID NO:123 is a clone designated herein as "DNA57037-1444".

Figure 124 shows the amino acid sequence (SEQ ID NO:124) derived from the coding sequence of SEQ ID NO:123 shown in Figure 123.

Figure 125 shows a nucleotide sequence (SEQ ID NO:125) of a native sequence PRO827 cDNA, wherein SEQ ID NO:125 is a clone designated herein as "DNA57039-1402".

10       Figure 126 shows the amino acid sequence (SEQ ID NO:126) derived from the coding sequence of SEQ ID NO:125 shown in Figure 125.

Figure 127 shows a nucleotide sequence (SEQ ID NO:127) of a native sequence PRO1075 cDNA, wherein SEQ ID NO:127 is a clone designated herein as "DNA57689-1385".

15       Figure 128 shows the amino acid sequence (SEQ ID NO:128) derived from the coding sequence of SEQ ID NO:127 shown in Figure 127.

Figure 129 shows a nucleotide sequence (SEQ ID NO:129) of a native sequence PRO1007 cDNA, wherein SEQ ID NO:129 is a clone designated herein as "DNA57690-1374".

Figure 130 shows the amino acid sequence (SEQ ID NO:130) derived from the coding sequence of SEQ ID NO:129 shown in Figure 129.

20       Figure 131 shows a nucleotide sequence (SEQ ID NO:131) of a native sequence PRO826 cDNA, wherein SEQ ID NO:131 is a clone designated herein as "DNA57694-1341".

Figure 132 shows the amino acid sequence (SEQ ID NO:132) derived from the coding sequence of SEQ ID NO:131 shown in Figure 131.

25       Figure 133 shows a nucleotide sequence (SEQ ID NO:133) of a native sequence PRO819 cDNA, wherein SEQ ID NO:132 is a clone designated herein as "DNA57695-1340".

Figure 134 shows the amino acid sequence (SEQ ID NO:134) derived from the coding sequence of SEQ ID NO:133 shown in Figure 133.

Figure 135 shows a nucleotide sequence (SEQ ID NO:135) of a native sequence PRO1006 cDNA, wherein SEQ ID NO:135 is a clone designated herein as "DNA57699-1412".

30       Figure 136 shows the amino acid sequence (SEQ ID NO:136) derived from the coding sequence of SEQ ID NO:135 shown in Figure 135.

Figure 137 shows a nucleotide sequence (SEQ ID NO:137) of a native sequence PRO982 cDNA, wherein SEQ ID NO:137 is a clone designated herein as "DNA57700-1408".

35       Figure 138 shows the amino acid sequence (SEQ ID NO:138) derived from the coding sequence of SEQ ID NO:137 shown in Figure 137.

Figure 139 shows a nucleotide sequence (SEQ ID NO:139) of a native sequence PRO1005 cDNA, wherein SEQ ID NO:139 is a clone designated herein as "DNA57708-1411".

Figure 140 shows the amino acid sequence (SEQ ID NO:140) derived from the coding sequence of SEQ ID NO:139 shown in Figure 139.

Figure 141 shows a nucleotide sequence (SEQ ID NO:141) of a native sequence PRO791 cDNA, wherein SEQ ID NO:141 is a clone designated herein as "DNA57838-1337".

5        Figure 142 shows the amino acid sequence (SEQ ID NO:142) derived from the coding sequence of SEQ ID NO:141 shown in Figure 141.

Figure 143 shows a nucleotide sequence (SEQ ID NO:143) of a native sequence PRO1071 cDNA, wherein SEQ ID NO:143 is a clone designated herein as "DNA58847-1383".

10       Figure 144 shows the amino acid sequence (SEQ ID NO:144) derived from the coding sequence of SEQ ID NO:143 shown in Figure 43.

Figure 145 shows a nucleotide sequence (SEQ ID NO:145) of a native sequence PRO1415 cDNA, wherein SEQ ID NO:145 is a clone designated herein as "DNA58852-1637".

Figure 146 shows the amino acid sequence (SEQ ID NO:146) derived from the coding sequence of SEQ ID NO:145 shown in Figure 145.

15       Figure 147 shows a nucleotide sequence (SEQ ID NO:147) of a native sequence PRO1054 cDNA, wherein SEQ ID NO:147 is a clone designated herein as "DNA58853-1423".

Figure 148 shows the amino acid sequence (SEQ ID NO:148) derived from the coding sequence of SEQ ID NO:147 shown in Figure 147.

20       Figure 149 shows a nucleotide sequence (SEQ ID NO:149) of a native sequence PRO1411 cDNA, wherein SEQ ID NO:149 is a clone designated herein as "DNA59212-1627".

Figure 150 shows the amino acid sequence (SEQ ID NO:150) derived from the coding sequence of SEQ ID NO:149 shown in Figure 149.

Figure 151 shows a nucleotide sequence (SEQ ID NO:151) of a native sequence PRO1184 cDNA, wherein SEQ ID NO:151 is a clone designated herein as "DNA59220-1514".

25       Figure 152 shows the amino acid sequence (SEQ ID NO:152) derived from the coding sequence of SEQ ID NO:151 shown in Figure 151.

Figure 153 shows a nucleotide sequence (SEQ ID NO:153) of a native sequence PRO1029 cDNA, wherein SEQ ID NO:153 is a clone designated herein as "DNA59493-1420".

30       Figure 154 shows the amino acid sequence (SEQ ID NO:154) derived from the coding sequence of SEQ ID NO:153 shown in Figure 153.

Figure 155 shows a nucleotide sequence (SEQ ID NO:155) of a native sequence PRO1139 cDNA, wherein SEQ ID NO:155 is a clone designated herein as "DNA59497-1496".

Figure 156 shows the amino acid sequence (SEQ ID NO:156) derived from the coding sequence of SEQ ID NO:155 shown in Figure 155.

35       Figure 157 shows a nucleotide sequence (SEQ ID NO:157) of a native sequence PRO1190 cDNA, wherein SEQ ID NO:157 is a clone designated herein as "DNA59586-1520".

Figure 158 shows the amino acid sequence (SEQ ID NO:158) derived from the coding sequence of SEQ

ID NO:157 shown in Figure 157.

Figure 159 shows a nucleotide sequence (SEQ ID NO:159) of a native sequence PRO1309 cDNA, wherein SEQ ID NO:159 is a clone designated herein as "DNA59588-1571".

5 Figure 160 shows the amino acid sequence (SEQ ID NO:160) derived from the coding sequence of SEQ ID NO:159 shown in Figure 159.

Figure 161 shows a nucleotide sequence (SEQ ID NO:161) of a native sequence PRO836 cDNA, wherein SEQ ID NO:161 is a clone designated herein as "DNA59620-1463".

Figure 162 shows the amino acid sequence (SEQ ID NO:162) derived from the coding sequence of SEQ ID NO:161 shown in Figure 161.

10 Figure 163 shows a nucleotide sequence (SEQ ID NO:163) of a native sequence PRO1025 cDNA, wherein SEQ ID NO:163 is a clone designated herein as "DNA59622-1334".

Figure 164 shows the amino acid sequence (SEQ ID NO:164) derived from the coding sequence of SEQ ID NO:163 shown in Figure 163.

15 Figure 165 shows a nucleotide sequence (SEQ ID NO:165) of a native sequence PRO1131 cDNA, wherein SEQ ID NO:165 is a clone designated herein as "DNA59777-1480".

Figure 166 shows the amino acid sequence (SEQ ID NO:166) derived from the coding sequence of SEQ ID NO:165 shown in Figure 165.

Figure 167 shows a nucleotide sequence (SEQ ID NO:167) of a native sequence PRO1182 cDNA, wherein SEQ ID NO:167 is a clone designated herein as "DNA59848-1512".

20 Figure 168 shows the amino acid sequence (SEQ ID NO:168) derived from the coding sequence of SEQ ID NO:167 shown in Figure 167.

Figure 169 shows a nucleotide sequence (SEQ ID NO:169) of a native sequence PRO1155 cDNA, wherein SEQ ID NO:169 is a clone designated herein as "DNA59849-1504".

25 Figure 170 shows the amino acid sequence (SEQ ID NO:170) derived from the coding sequence of SEQ ID NO:169 shown in Figure 169.

Figure 171 shows a nucleotide sequence (SEQ ID NO:171) of a native sequence PRO1186 cDNA, wherein SEQ ID NO:171 is a clone designated herein as "DNA60621-1516".

Figure 172 shows the amino acid sequence (SEQ ID NO:172) derived from the coding sequence of SEQ ID NO:171 shown in Figure 171.

30 Figure 173 shows a nucleotide sequence (SEQ ID NO:173) of a native sequence PRO1198 cDNA, wherein SEQ ID NO:173 is a clone designated herein as "DNA60622-1525".

Figure 174 shows the amino acid sequence (SEQ ID NO:174) derived from the coding sequence of SEQ ID NO:173 shown in Figure 173.

35 Figure 175 shows a nucleotide sequence (SEQ ID NO:175) of a native sequence PRO1265 cDNA, wherein SEQ ID NO:175 is a clone designated herein as "DNA60764-1533".

Figure 176 shows the amino acid sequence (SEQ ID NO:176) derived from the coding sequence of SEQ ID NO:175 shown in Figure 175.

Figure 177 shows a nucleotide sequence (SEQ ID NO:177) of a native sequence PRO1361 cDNA, wherein SEQ ID NO:177 is a clone designated herein as "DNA60783-1611".

Figure 178 shows the amino acid sequence (SEQ ID NO:178) derived from the coding sequence of SEQ ID NO:177 shown in Figure 177.

5        Figure 179 shows a nucleotide sequence (SEQ ID NO:179) of a native sequence PRO1287 cDNA, wherein SEQ ID NO:179 is a clone designated herein as "DNA61755-1554".

Figure 180 shows the amino acid sequence (SEQ ID NO:180) derived from the coding sequence of SEQ ID NO:179 shown in Figure 179.

10       Figure 181 shows a nucleotide sequence (SEQ ID NO:181) of a native sequence PRO1308 cDNA, wherein SEQ ID NO:181 is a clone designated herein as "DNA62306-1570".

Figure 182 shows the amino acid sequence (SEQ ID NO:182) derived from the coding sequence of SEQ ID NO:181 shown in Figure 181.

Figure 183 shows a nucleotide sequence (SEQ ID NO:183) of a native sequence PRO4313 cDNA, wherein SEQ ID NO:183 is a clone designated herein as "DNA62312-2558".

15       Figure 184 shows the amino acid sequence (SEQ ID NO:184) derived from the coding sequence of SEQ ID NO:183 shown in Figure 183.

Figure 185 shows a nucleotide sequence (SEQ ID NO:185) of a native sequence PRO1192 cDNA, wherein SEQ ID NO:185 is a clone designated herein as "DNA62814-1521".

20       Figure 186 shows the amino acid sequence (SEQ ID NO:186) derived from the coding sequence of SEQ ID NO:185 shown in Figure 185.

Figure 187 shows a nucleotide sequence (SEQ ID NO:187) of a native sequence PRO1160 cDNA, wherein SEQ ID NO:187 is a clone designated herein as "DNA62872-1509".

Figure 188 shows the amino acid sequence (SEQ ID NO:188) derived from the coding sequence of SEQ ID NO:187 shown in Figure 187.

25       Figure 189 shows a nucleotide sequence (SEQ ID NO:189) of a native sequence PRO1244 cDNA, wherein SEQ ID NO:189 is a clone designated herein as "DNA64883-1526".

Figure 190 shows the amino acid sequence (SEQ ID NO:190) derived from the coding sequence of SEQ ID NO:189 shown in Figure 189.

30       Figure 191 shows a nucleotide sequence (SEQ ID NO:191) of a native sequence PRO1356 cDNA, wherein SEQ ID NO:191 is a clone designated herein as "DNA64886-1601".

Figure 192 shows the amino acid sequence (SEQ ID NO:192) derived from the coding sequence of SEQ ID NO:191 shown in Figure 191.

Figure 193 shows a nucleotide sequence (SEQ ID NO:193) of a native sequence PRO1274 cDNA, wherein SEQ ID NO:193 is a clone designated herein as "DNA64889-1541".

35       Figure 194 shows the amino acid sequence (SEQ ID NO:194) derived from the coding sequence of SEQ ID NO:193 shown in Figure 193.

Figure 195 shows a nucleotide sequence (SEQ ID NO:195) of a native sequence PRO1272 cDNA, wherein



SEQ ID NO:195 is a clone designated herein as "DNA64896-1539".

Figure 196 shows the amino acid sequence (SEQ ID NO:196) derived from the coding sequence of SEQ ID NO:195 shown in Figure 195.

5 Figure 197 shows a nucleotide sequence (SEQ ID NO:197) of a native sequence PRO1412 cDNA, wherein SEQ ID NO:197 is a clone designated herein as "DNA64897-1628".

Figure 198 shows the amino acid sequence (SEQ ID NO:198) derived from the coding sequence of SEQ ID NO:197 shown in Figure 197.

Figure 199 shows a nucleotide sequence (SEQ ID NO:199) of a native sequence PRO1286 cDNA, wherein SEQ ID NO:199 is a clone designated herein as "DNA64903-1553".

10 Figure 200 shows the amino acid sequence (SEQ ID NO:200) derived from the coding sequence of SEQ ID NO:199 shown in Figure 199.

Figure 201 shows a nucleotide sequence (SEQ ID NO:201) of a native sequence PRO1347 cDNA, wherein SEQ ID NO:201 is a clone designated herein as "DNA64950-1590".

15 Figure 202 shows the amino acid sequence (SEQ ID NO:202) derived from the coding sequence of SEQ ID NO:201 shown in Figure 201.

Figure 203 shows a nucleotide sequence (SEQ ID NO:203) of a native sequence PRO1273 cDNA, wherein SEQ ID NO:203 is a clone designated herein as "DNA65402-1540".

Figure 204 shows the amino acid sequence (SEQ ID NO:204) derived from the coding sequence of SEQ ID NO:203 shown in Figure 203.

20 Figure 205 shows a nucleotide sequence (SEQ ID NO:205) of a native sequence PRO1283 cDNA, wherein SEQ ID NO:205 is a clone designated herein as "DNA65404-1551".

Figure 206 shows the amino acid sequence (SEQ ID NO:206) derived from the coding sequence of SEQ ID NO:205 shown in Figure 205.

25 Figure 207 shows a nucleotide sequence (SEQ ID NO:207) of a native sequence PRO1279 cDNA, wherein SEQ ID NO:207 is a clone designated herein as "DNA65405-1547".

Figure 208 shows the amino acid sequence (SEQ ID NO:208) derived from the coding sequence of SEQ ID NO:207 shown in Figure 207.

Figure 209 shows a nucleotide sequence (SEQ ID NO:209) of a native sequence PRO1306 cDNA, wherein SEQ ID NO:209 is a clone designated herein as "DNA65410-1569".

30 Figure 210 shows the amino acid sequence (SEQ ID NO:210) derived from the coding sequence of SEQ ID NO:209 shown in Figure 209.

Figure 211 shows a nucleotide sequence (SEQ ID NO:211) of a native sequence PRO1195 cDNA, wherein SEQ ID NO:211 is a clone designated herein as "DNA65412-1523".

35 Figure 212 shows the amino acid sequence (SEQ ID NO:212) derived from the coding sequence of SEQ ID NO:211 shown in Figure 211.

Figure 213 shows a nucleotide sequence (SEQ ID NO:213) of a native sequence PRO4995 cDNA, wherein SEQ ID NO:213 is a clone designated herein as "DNA66307-2661".

Figure 214 shows the amino acid sequence (SEQ ID NO:214) derived from the coding sequence of SEQ ID NO:213 shown in Figure 213.

Figure 215 shows a nucleotide sequence (SEQ ID NO:215) of a native sequence PRO1382 cDNA, wherein SEQ ID NO:215 is a clone designated herein as "DNA66526-1616".

5        Figure 216 shows the amino acid sequence (SEQ ID NO:216) derived from the coding sequence of SEQ ID NO:215 shown in Figure 215.

Figure 217 shows a nucleotide sequence (SEQ ID NO:217) of a native sequence PRO1325 cDNA, wherein SEQ ID NO:217 is a clone designated herein as "DNA66659-1593".

10       Figure 218 shows the amino acid sequence (SEQ ID NO:218) derived from the coding sequence of SEQ ID NO:217 shown in Figure 217.

Figure 219 shows a nucleotide sequence (SEQ ID NO:219) of a native sequence PRO1329 cDNA, wherein SEQ ID NO:219 is a clone designated herein as "DNA66660-1585".

Figure 220 shows the amino acid sequence (SEQ ID NO:220) derived from the coding sequence of SEQ ID NO:219 shown in Figure 219.

15       Figure 221 shows a nucleotide sequence (SEQ ID NO:221) of a native sequence PRO1338 cDNA, wherein SEQ ID NO:221 is a clone designated herein as "DNA66667-1596".

Figure 222 shows the amino acid sequence (SEQ ID NO:222) derived from the coding sequence of SEQ ID NO:221 shown in Figure 221.

20       Figure 223 shows a nucleotide sequence (SEQ ID NO:223) of a native sequence PRO1337 cDNA, wherein SEQ ID NO:223 is a clone designated herein as "DNA66672-1586".

Figure 224 shows the amino acid sequence (SEQ ID NO:224) derived from the coding sequence of SEQ ID NO:223 shown in Figure 223.

Figure 225 shows a nucleotide sequence (SEQ ID NO:225) of a native sequence PRO1343 cDNA, wherein SEQ ID NO:225 is a clone designated herein as "DNA66675-1587".

25       Figure 226 shows the amino acid sequence (SEQ ID NO:226) derived from the coding sequence of SEQ ID NO:225 shown in Figure 225.

Figure 227 shows a nucleotide sequence (SEQ ID NO:227) of a native sequence PRO1376 cDNA, wherein SEQ ID NO:227 is a clone designated herein as "DNA67300-1605".

30       Figure 228 shows the amino acid sequence (SEQ ID NO:228) derived from the coding sequence of SEQ ID NO:227 shown in Figure 227.

Figure 229 shows a nucleotide sequence (SEQ ID NO:229) of a native sequence PRO1434 cDNA, wherein SEQ ID NO:229 is a clone designated herein as "DNA68818-2536".

Figure 230 shows the amino acid sequence (SEQ ID NO:230) derived from the coding sequence of SEQ ID NO:229 shown in Figure 229.

35       Figure 231 shows a nucleotide sequence (SEQ ID NO:231) of a native sequence PRO3579 cDNA, wherein SEQ ID NO:231 is a clone designated herein as "DNA68862-2546".

Figure 232 shows the amino acid sequence (SEQ ID NO:232) derived from the coding sequence of SEQ

ID NO:231 shown in Figure 231.

Figure 233 shows a nucleotide sequence (SEQ ID NO:233) of a native sequence PRO1387 cDNA, wherein SEQ ID NO:233 is a clone designated herein as "DNA68872-1620".

5 Figure 234 shows the amino acid sequence (SEQ ID NO:234) derived from the coding sequence of SEQ ID NO:233 shown in Figure 233.

Figure 235 shows a nucleotide sequence (SEQ ID NO:235) of a native sequence PRO1419 cDNA, wherein SEQ ID NO:235 is a clone designated herein as "DNA71290-1630".

Figure 236 shows the amino acid sequence (SEQ ID NO:236) derived from the coding sequence of SEQ ID NO:235 shown in Figure 235.

10 Figure 237 shows a nucleotide sequence (SEQ ID NO:237) of a native sequence PRO1488 cDNA, wherein SEQ ID NO:237 is a clone designated herein as "DNA73736-1657".

Figure 238 shows the amino acid sequence (SEQ ID NO:238) derived from the coding sequence of SEQ ID NO:237 shown in Figure 237.

15 Figure 239 shows a nucleotide sequence (SEQ ID NO:239) of a native sequence PRO1474 cDNA, wherein SEQ ID NO:239 is a clone designated herein as "DNA73739-1645".

Figure 240 shows the amino acid sequence (SEQ ID NO:240) derived from the coding sequence of SEQ ID NO:239 shown in Figure 239.

Figure 241 shows a nucleotide sequence (SEQ ID NO:241) of a native sequence PRO1508 cDNA, wherein SEQ ID NO:241 is a clone designated herein as "DNA73742-1662".

20 Figure 242 shows the amino acid sequence (SEQ ID NO:242) derived from the coding sequence of SEQ ID NO:241 shown in Figure 241.

Figure 243 shows a nucleotide sequence (SEQ ID NO:243) of a native sequence PRO1754 cDNA, wherein SEQ ID NO:243 is a clone designated herein as "DNA76385-1692".

25 Figure 244 shows the amino acid sequence (SEQ ID NO:244) derived from the coding sequence of SEQ ID NO:243 shown in Figure 243.

Figure 245 shows a nucleotide sequence (SEQ ID NO:245) of a native sequence PRO1550 cDNA, wherein SEQ ID NO:245 is a clone designated herein as "DNA76393-1664".

Figure 246 shows the amino acid sequence (SEQ ID NO:246) derived from the coding sequence of SEQ ID NO:245 shown in Figure 245.

30 Figure 247 shows a nucleotide sequence (SEQ ID NO:247) of a native sequence PRO1758 cDNA, wherein SEQ ID NO:247 is a clone designated herein as "DNA76399-1700".

Figure 248 shows the amino acid sequence (SEQ ID NO:248) derived from the coding sequence of SEQ ID NO:247 shown in Figure 247.

35 Figure 249 shows a nucleotide sequence (SEQ ID NO:249) of a native sequence PRO1917 cDNA, wherein SEQ ID NO:249 is a clone designated herein as "DNA76400-2528".

Figure 250 shows the amino acid sequence (SEQ ID NO:250) derived from the coding sequence of SEQ ID NO:249 shown in Figure 249.

Figure 251 shows a nucleotide sequence (SEQ ID NO:251) of a native sequence PRO1787 cDNA, wherein SEQ ID NO:251 is a clone designated herein as "DNA76510-2504".

Figure 252 shows the amino acid sequence (SEQ ID NO:252) derived from the coding sequence of SEQ ID NO:251 shown in Figure 251.

5        Figure 253 shows a nucleotide sequence (SEQ ID NO:253) of a native sequence PRO1556 cDNA, wherein SEQ ID NO:253 is a clone designated herein as "DNA76529-1666".

Figure 254 shows the amino acid sequence (SEQ ID NO:254) derived from the coding sequence of SEQ ID NO:253 shown in Figure 253.

10       Figure 255 shows a nucleotide sequence (SEQ ID NO:255) of a native sequence PRO1760 cDNA, wherein SEQ ID NO:255 is a clone designated herein as "DNA76532-1702".

Figure 256 shows the amino acid sequence (SEQ ID NO:256) derived from the coding sequence of SEQ ID NO:255 shown in Figure 255.

Figure 257 shows a nucleotide sequence (SEQ ID NO:257) of a native sequence PRO1567 cDNA, wherein SEQ ID NO:257 is a clone designated herein as "DNA76541-1675".

15       Figure 258 shows the amino acid sequence (SEQ ID NO:258) derived from the coding sequence of SEQ ID NO:257 shown in Figure 257.

Figure 259 shows a nucleotide sequence (SEQ ID NO:259) of a native sequence PRO1600 cDNA, wherein SEQ ID NO:259 is a clone designated herein as "DNA77503-1686".

20       Figure 260 shows the amino acid sequence (SEQ ID NO:260) derived from the coding sequence of SEQ ID NO:259 shown in Figure 259.

Figure 261 shows a nucleotide sequence (SEQ ID NO:261) of a native sequence PRO1868 cDNA, wherein SEQ ID NO:261 is a clone designated herein as "DNA77624-2515".

Figure 262 shows the amino acid sequence (SEQ ID NO:262) derived from the coding sequence of SEQ ID NO:261 shown in Figure 261.

25       Figure 263 shows a nucleotide sequence (SEQ ID NO:263) of a native sequence PRO1890 cDNA, wherein SEQ ID NO:263 is a clone designated herein as "DNA79230-2525".

Figure 264 shows the amino acid sequence (SEQ ID NO:264) derived from the coding sequence of SEQ ID NO:263 shown in Figure 263.

30       Figure 265 shows a nucleotide sequence (SEQ ID NO:265) of a native sequence PRO1887 cDNA, wherein SEQ ID NO:265 is a clone designated herein as "DNA79862-2522".

Figure 266 shows the amino acid sequence (SEQ ID NO:265) derived from the coding sequence of SEQ ID NO:265 shown in Figure 265.

Figure 267 shows a nucleotide sequence (SEQ ID NO:267) of a native sequence PRO4353 cDNA, wherein SEQ ID NO:267 is a clone designated herein as "DNA80145-2594".

35       Figure 268 shows the amino acid sequence (SEQ ID NO:268) derived from the coding sequence of SEQ ID NO:267 shown in Figure 267.

Figure 269 shows a nucleotide sequence (SEQ ID NO:269) of a native sequence PRO1782 cDNA, wherein

SEQ ID NO:269 is a clone designated herein as "DNA80899-2501".

Figure 270 shows the amino acid sequence (SEQ ID NO:270) derived from the coding sequence of SEQ ID NO:269 shown in Figure 269.

5 Figure 271 shows a nucleotide sequence (SEQ ID NO:271) of a native sequence PRO1928 cDNA, wherein SEQ ID NO:271 is a clone designated herein as "DNA81754-2532".

Figure 272 shows the amino acid sequence (SEQ ID NO:272) derived from the coding sequence of SEQ ID NO:271 shown in Figure 271.

Figure 273 shows a nucleotide sequence (SEQ ID NO:273) of a native sequence PRO1865 cDNA, wherein SEQ ID NO:273 is a clone designated herein as "DNA81757-2512".

10 Figure 274 shows the amino acid sequence (SEQ ID NO:274) derived from the coding sequence of SEQ ID NO:273 shown in Figure 273.

Figure 275 shows a nucleotide sequence (SEQ ID NO:275) of a native sequence PRO4341 cDNA, wherein SEQ ID NO:275 is a clone designated herein as "DNA81761-2583".

15 Figure 276 shows the amino acid sequence (SEQ ID NO:276) derived from the coding sequence of SEQ ID NO:275 shown in Figure 275.

Figure 277 shows a nucleotide sequence (SEQ ID NO:277) of a native sequence PRO6714 cDNA, wherein SEQ ID NO:277 is a clone designated herein as "DNA82358-2738".

Figure 278 shows the amino acid sequence (SEQ ID NO:278) derived from the coding sequence of SEQ ID NO:277 shown in Figure 277.

20 Figure 279 shows a nucleotide sequence (SEQ ID NO:279) of a native sequence PRO5723 cDNA, wherein SEQ ID NO:279 is a clone designated herein as "DNA82361".

Figure 280 shows the amino acid sequence (SEQ ID NO:280) derived from the coding sequence of SEQ ID NO:279 shown in Figure 279.

25 Figure 281 shows a nucleotide sequence (SEQ ID NO:281) of a native sequence PRO3438 cDNA, wherein SEQ ID NO:281 is a clone designated herein as "DNA82364-2538".

Figure 282 shows the amino acid sequence (SEQ ID NO:282) derived from the coding sequence of SEQ ID NO:281 shown in Figure 281.

Figure 283 shows a nucleotide sequence (SEQ ID NO:283) of a native sequence PRO6071 cDNA, wherein SEQ ID NO:283 is a clone designated herein as "DNA82403-2959".

30 Figure 284 shows the amino acid sequence (SEQ ID NO:284) derived from the coding sequence of SEQ ID NO:283 shown in Figure 283.

Figure 285 shows a nucleotide sequence (SEQ ID NO:285) of a native sequence PRO1801 cDNA, wherein SEQ ID NO:285 is a clone designated herein as "DNA83500-2506".

35 Figure 286 shows the amino acid sequence (SEQ ID NO:286) derived from the coding sequence of SEQ ID NO:285 shown in Figure 285.

Figure 287 shows a nucleotide sequence (SEQ ID NO:287) of a native sequence PRO4324 cDNA, wherein SEQ ID NO:287 is a clone designated herein as "DNA83560-2569".

Figure 288 shows the amino acid sequence (SEQ ID NO:288) derived from the coding sequence of SEQ ID NO:287 shown in Figure 287.

Figure 289 shows a nucleotide sequence (SEQ ID NO:289) of a native sequence PRO4333 cDNA, wherein SEQ ID NO:289 is a clone designated herein as "DNA84210-2576".

5        Figure 290 shows the amino acid sequence (SEQ ID NO:290) derived from the coding sequence of SEQ ID NO:289 shown in Figure 289.

Figure 291 shows a nucleotide sequence (SEQ ID NO:291) of a native sequence PRO4405 cDNA, wherein SEQ ID NO:291 is a clone designated herein as "DNA84920-2614".

10       Figure 292 shows the amino acid sequence (SEQ ID NO:292) derived from the coding sequence of SEQ ID NO:291 shown in Figure 291.

Figure 293 shows a nucleotide sequence (SEQ ID NO:293) of a native sequence PRO4356 cDNA, wherein SEQ ID NO:293 is a clone designated herein as "DNA86576-2595".

Figure 294 shows the amino acid sequence (SEQ ID NO:294) derived from the coding sequence of SEQ ID NO:293 shown in Figure 293.

15       Figure 295 shows a nucleotide sequence (SEQ ID NO:295) of a native sequence PRO3444 cDNA, wherein SEQ ID NO:295 is a clone designated herein as "DNA87997".

Figure 296 shows the amino acid sequence (SEQ ID NO:296) derived from the coding sequence of SEQ ID NO:295 shown in Figure 295.

20       Figure 297 shows a nucleotide sequence (SEQ ID NO:297) of a native sequence PRO4302 cDNA, wherein SEQ ID NO:297 is a clone designated herein as "DNA92218-2554".

Figure 298 shows the amino acid sequence (SEQ ID NO:298) derived from the coding sequence of SEQ ID NO:297 shown in Figure 297.

Figure 299 shows a nucleotide sequence (SEQ ID NO:299) of a native sequence PRO4371 cDNA, wherein SEQ ID NO:299 is a clone designated herein as "DNA92233-2599".

25       Figure 300 shows the amino acid sequence (SEQ ID NO:300) derived from the coding sequence of SEQ ID NO:299 shown in Figure 299.

Figure 301 shows a nucleotide sequence (SEQ ID NO:301) of a native sequence PRO4354 cDNA, wherein SEQ ID NO:301 is a clone designated herein as "DNA92256-2596".

30       Figure 302 shows the amino acid sequence (SEQ ID NO:302) derived from the coding sequence of SEQ ID NO:301 shown in Figure 301.

Figure 303 shows a nucleotide sequence (SEQ ID NO:303) of a native sequence PRO5725 cDNA, wherein SEQ ID NO:303 is a clone designated herein as "DNA92265-2669".

Figure 304 shows the amino acid sequence (SEQ ID NO:304) derived from the coding sequence of SEQ ID NO:303 shown in Figure 303.

35       Figure 305 shows a nucleotide sequence (SEQ ID NO:305) of a native sequence PRO4408 cDNA, wherein SEQ ID NO:305 is a clone designated herein as "DNA92274-2617".

Figure 306 shows the amino acid sequence (SEQ ID NO:306) derived from the coding sequence of SEQ

ID NO:305 shown in Figure 305.

Figure 307 shows a nucleotide sequence (SEQ ID NO:307) of a native sequence PRO9940 cDNA, wherein SEQ ID NO:307 is a clone designated herein as "DNA92282".

5 Figure 308 shows the amino acid sequence (SEQ ID NO:308) derived from the coding sequence of SEQ ID NO:307 shown in Figure 307.

Figure 309 shows a nucleotide sequence (SEQ ID NO:309) of a native sequence PRO5737 cDNA, wherein SEQ ID NO:309 is a clone designated herein as "DNA92929-2534-1".

Figure 310 shows the amino acid sequence (SEQ ID NO:310) derived from the coding sequence of SEQ ID NO:309 shown in Figure 309.

10 Figure 311 shows a nucleotide sequence (SEQ ID NO:311) of a native sequence PRO4425 cDNA, wherein SEQ ID NO:311 is a clone designated herein as "DNA93011-2637".

Figure 312 shows the amino acid sequence (SEQ ID NO:312) derived from the coding sequence of SEQ ID NO:311 shown in Figure 311.

15 Figure 313 shows a nucleotide sequence (SEQ ID NO:313) of a native sequence PRO4345 cDNA, wherein SEQ ID NO:313 is a clone designated herein as "DNA94854-2586".

Figure 314 shows the amino acid sequence (SEQ ID NO:314) derived from the coding sequence of SEQ ID NO:313 shown in Figure 313.

Figure 315 shows a nucleotide sequence (SEQ ID NO:315) of a native sequence PRO4342 cDNA, wherein SEQ ID NO:315 is a clone designated herein as "DNA96787-2534-1".

20 Figure 316 shows the amino acid sequence (SEQ ID NO:316) derived from the coding sequence of SEQ ID NO:315 shown in Figure 315.

Figure 317 shows a nucleotide sequence (SEQ ID NO:317) of a native sequence PRO3562 cDNA, wherein SEQ ID NO:317 is a clone designated herein as "DNA96791".

25 Figure 318 shows the amino acid sequence (SEQ ID NO:318) derived from the coding sequence of SEQ ID NO:317 shown in Figure 317.

Figure 319 shows a nucleotide sequence (SEQ ID NO:319) of a native sequence PRO4422 cDNA, wherein SEQ ID NO:319 is a clone designated herein as "DNA96867-2620".

Figure 320 shows the amino acid sequence (SEQ ID NO:320) derived from the coding sequence of SEQ ID NO:319 shown in Figure 319.

30 Figure 321 shows a nucleotide sequence (SEQ ID NO:321) of a native sequence PRO5776 cDNA, wherein SEQ ID NO:321 is a clone designated herein as "DNA96872-2674".

Figure 322 shows the amino acid sequence (SEQ ID NO:322) derived from the coding sequence of SEQ ID NO:321 shown in Figure 321.

35 Figure 323 shows a nucleotide sequence (SEQ ID NO:323) of a native sequence PRO4430 cDNA, wherein SEQ ID NO:323 is a clone designated herein as "DNA96878-2626".

Figure 324 shows the amino acid sequence (SEQ ID NO:324) derived from the coding sequence of SEQ ID NO:323 shown in Figure 323.

Figure 325 shows a nucleotide sequence (SEQ ID NO:325) of a native sequence PRO4499 cDNA, wherein SEQ ID NO:325 is a clone designated herein as "DNA96889-2641".

Figure 326 shows the amino acid sequence (SEQ ID NO:326) derived from the coding sequence of SEQ ID NO:325 shown in Figure 325.

5        Figure 327 shows a nucleotide sequence (SEQ ID NO:327) of a native sequence PRO4503 cDNA, wherein SEQ ID NO:327 is a clone designated herein as "DNA100312-2645".

Figure 328 shows the amino acid sequence (SEQ ID NO:328) derived from the coding sequence of SEQ ID NO:327 shown in Figure 327.

10       Figure 329 shows a nucleotide sequence (SEQ ID NO:329) of a native sequence PRO10008 cDNA, wherein SEQ ID NO:329 is a clone designated herein as "DNA101921".

Figure 330 shows the amino acid sequence (SEQ ID NO:330) derived from the coding sequence of SEQ ID NO:329 shown in Figure 329.

Figure 331 shows a nucleotide sequence (SEQ ID NO:331) of a native sequence PRO5730 cDNA, wherein SEQ ID NO:331 is a clone designated herein as "DNA101926".

15       Figure 332 shows the amino acid sequence (SEQ ID NO:332) derived from the coding sequence of SEQ ID NO:331 shown in Figure 331.

Figure 333 shows a nucleotide sequence (SEQ ID NO:333) of a native sequence PRO6008 cDNA, wherein SEQ ID NO:333 is a clone designated herein as "DNA102844".

20       Figure 334 shows the amino acid sequence (SEQ ID NO:334) derived from the coding sequence of SEQ ID NO:333 shown in Figure 333.

Figure 335 shows a nucleotide sequence (SEQ ID NO:335) of a native sequence PRO4527 cDNA, wherein SEQ ID NO:335 is a clone designated herein as "DNA103197".

Figure 336 shows the amino acid sequence (SEQ ID NO:336) derived from the coding sequence of SEQ ID NO:335 shown in Figure 335.

25       Figure 337 shows a nucleotide sequence (SEQ ID NO:337) of a native sequence PRO4538 cDNA, wherein SEQ ID NO:337 is a clone designated herein as "DNA103208".

Figure 338 shows the amino acid sequence (SEQ ID NO:338) derived from the coding sequence of SEQ ID NO:337 shown in Figure 337.

30       Figure 339 shows a nucleotide sequence (SEQ ID NO:339) of a native sequence PRO4553 cDNA, wherein SEQ ID NO:339 is a clone designated herein as "DNA103223".

Figure 340 shows the amino acid sequence (SEQ ID NO:340) derived from the coding sequence of SEQ ID NO:339 shown in Figure 339.

Figure 341 shows a nucleotide sequence (SEQ ID NO:341) of a native sequence PRO6006 cDNA, wherein SEQ ID NO:341 is a clone designated herein as "DNA105782-2693".

35       Figure 342 shows the amino acid sequence (SEQ ID NO:342) derived from the coding sequence of SEQ ID NO:341 shown in Figure 341.

Figure 343 shows a nucleotide sequence (SEQ ID NO:343) of a native sequence PRO6029 cDNA, wherein



SEQ ID NO:343 is a clone designated herein as "DNA105849-2704".

Figure 344 shows the amino acid sequence (SEQ ID NO:344) derived from the coding sequence of SEQ ID NO:343 shown in Figure 343.

5 Figure 345 shows a nucleotide sequence (SEQ ID NO:345) of a native sequence PRO9821 cDNA, wherein SEQ ID NO:345 is a clone designated herein as "DNA108725-2766".

Figure 346 shows the amino acid sequence (SEQ ID NO:346) derived from the coding sequence of SEQ ID NO:345 shown in Figure 345.

Figure 347 shows a nucleotide sequence (SEQ ID NO:347) of a native sequence PRO9820 cDNA, wherein SEQ ID NO:347 is a clone designated herein as "DNA108769-2765".

10 Figure 348 shows the amino acid sequence (SEQ ID NO:348) derived from the coding sequence of SEQ ID NO:347 shown in Figure 347.

Figure 349 shows a nucleotide sequence (SEQ ID NO:349) of a native sequence PRO9771 cDNA, wherein SEQ ID NO:349 is a clone designated herein as "DNA119498-2965".

15 Figure 350 shows the amino acid sequence (SEQ ID NO:350) derived from the coding sequence of SEQ ID NO:349 shown in Figure 349.

Figure 351 shows a nucleotide sequence (SEQ ID NO:351) of a native sequence PRO7436 cDNA, wherein SEQ ID NO:351 is a clone designated herein as "DNA119535-2756".

Figure 352 shows the amino acid sequence (SEQ ID NO:352) derived from the coding sequence of SEQ ID NO:351 shown in Figure 351.

20 Figure 353 shows a nucleotide sequence (SEQ ID NO:353) of a native sequence PRO10096 cDNA, wherein SEQ ID NO:353 is a clone designated herein as "DNA125185-2806".

Figure 354 shows the amino acid sequence (SEQ ID NO:354) derived from the coding sequence of SEQ ID NO:353 shown in Figure 353.

25 Figure 355 shows a nucleotide sequence (SEQ ID NO:355) of a native sequence PRO19670 cDNA, wherein SEQ ID NO:355 is a clone designated herein as "DNA131639-2874".

Figure 356 shows the amino acid sequence (SEQ ID NO:356) derived from the coding sequence of SEQ ID NO:355 shown in Figure 355.

Figure 357 shows a nucleotide sequence (SEQ ID NO:357) of a native sequence PRO20044 cDNA, wherein SEQ ID NO:357 is a clone designated herein as "DNA139623-2893".

30 Figure 358 shows the amino acid sequence (SEQ ID NO:358) derived from the coding sequence of SEQ ID NO:357 shown in Figure 357.

Figure 359 shows a nucleotide sequence (SEQ ID NO:359) of a native sequence PRO9873 cDNA, wherein SEQ ID NO:359 is a clone designated herein as "DNA143076-2787".

35 Figure 360 shows the amino acid sequence (SEQ ID NO:360) derived from the coding sequence of SEQ ID NO:359 shown in Figure 359.

Figure 361 shows a nucleotide sequence (SEQ ID NO:361) of a native sequence PRO21366 cDNA, wherein SEQ ID NO:361 is a clone designated herein as "DNA143276-2975".

Figure 362 shows the amino acid sequence (SEQ ID NO:362) derived from the coding sequence of SEQ ID NO:361 shown in Figure 361.

Figure 363 shows a nucleotide sequence (SEQ ID NO:363) of a native sequence PRO20040 cDNA, wherein SEQ ID NO:363 is a clone designated herein as "DNA164625-2890".

5        Figure 364 shows the amino acid sequence (SEQ ID NO:364) derived from the coding sequence of SEQ ID NO:363 shown in Figure 363.

Figure 365 shows a nucleotide sequence (SEQ ID NO:365) of a native sequence PRO21184 cDNA, wherein SEQ ID NO:365 is a clone designated herein as "DNA167678-2963".

10       Figure 366 shows the amino acid sequence (SEQ ID NO:366) derived from the coding sequence of SEQ ID NO:365 shown in Figure 365.

Figure 367 shows a nucleotide sequence (SEQ ID NO:367) of a native sequence PRO21055 cDNA, wherein SEQ ID NO:367 is a clone designated herein as "DNA170021-2923".

Figure 368 shows the amino acid sequence (SEQ ID NO:368) derived from the coding sequence of SEQ ID NO:367 shown in Figure 367.

15       Figure 369 shows a nucleotide sequence (SEQ ID NO:369) of a native sequence PRO28631 cDNA, wherein SEQ ID NO:369 is a clone designated herein as "DNA170212-3000".

Figure 370 shows the amino acid sequence (SEQ ID NO:370) derived from the coding sequence of SEQ ID NO:369 shown in Figure 369.

20       Figure 371 shows a nucleotide sequence (SEQ ID NO:371) of a native sequence PRO21384 cDNA, wherein SEQ ID NO:371 is a clone designated herein as "DNA177313-2982".

Figure 372 shows the amino acid sequence (SEQ ID NO:372) derived from the coding sequence of SEQ ID NO:371 shown in Figure 371.

Figure 373 shows a nucleotide sequence (SEQ ID NO:373) of a native sequence PRO1449 cDNA, wherein SEQ ID NO:373 is a clone designated herein as "DNA64908-1163-1".

25       Figure 374 shows the amino acid sequence (SEQ ID NO:374) derived from the coding sequence of SEQ ID NO:373 shown in Figure 373.

30       Figure 375 shows wholemount in situ hybridization results on mouse embryos using a mouse orthologue of PRO1449 which has about 78% amino acid identity with PRO1449. The results show that PRO1449 orthologue is expressed in the developing vasculature. The cross-section further shows expression in endothelial cells and progenitors of endothelial cells.

Figure 376 shows that a PRO1449 orthologue having about 78% amino acid identity with PRO1449 is expressed in vasculature of many inflamed and diseased tissues, but is very low, or lacking, in normal adult vessels.

Figure 377 shows that a PRO1449 orthologue having about 78% amino acid identity with PRO1449 induces ectopic vessels in the eyes of chicken embryos.

35

## 5. Detailed Description of the Invention

### 5.1. Definitions

The phrases "cardiovascular, endothelial and angiogenic disorder", "cardiovascular, endothelial and angiogenic dysfunction", "cardiovascular, endothelial or angiogenic disorder" and "cardiovascular, endothelial or angiogenic dysfunction" are used interchangeably and refer in part to systemic disorders that affect vessels, such as diabetes mellitus, as well as diseases of the vessels themselves, such as of the arteries, capillaries, veins, and/or lymphatics. This would include indications that stimulate angiogenesis and/or cardiovascularization, and those that inhibit angiogenesis and/or cardiovascularization. Such disorders include, for example, arterial disease, such as atherosclerosis, hypertension, inflammatory vasculitides, Reynaud's disease and Reynaud's phenomenon, aneurysms, and arterial restenosis; venous and lymphatic disorders such as thrombophlebitis, lymphangitis, and lymphedema; and other vascular disorders such as peripheral vascular disease, cancer such as vascular tumors, *e.g.*, hemangioma (capillary and cavernous), glomus tumors, telangiectasia, bacillary angiomatosis, hemangioendothelioma, angiosarcoma, haemangiopericytoma, Kaposi's sarcoma, lymphangioma, and lymphangiosarcoma, tumor angiogenesis, trauma such as wounds, burns, and other injured tissue, implant fixation, scarring, ischemia reperfusion injury, rheumatoid arthritis, cerebrovascular disease, renal diseases such as acute renal failure, and osteoporosis. This would also include angina, myocardial infarctions such as acute myocardial infarctions, cardiac hypertrophy, and heart failure such as CHF.

"Hypertrophy", as used herein, is defined as an increase in mass of an organ or structure independent of natural growth that does not involve tumor formation. Hypertrophy of an organ or tissue is due either to an increase in the mass of the individual cells (true hypertrophy), or to an increase in the number of cells making up the tissue (hyperplasia), or both. Certain organs, such as the heart, lose the ability to divide shortly after birth. Accordingly, "cardiac hypertrophy" is defined as an increase in mass of the heart, which, in adults, is characterized by an increase in myocyte cell size and contractile protein content without concomitant cell division. The character of the stress responsible for inciting the hypertrophy, (*e.g.*, increased preload, increased afterload, loss of myocytes, as in myocardial infarction, or primary depression of contractility), appears to play a critical role in determining the nature of the response. The early stage of cardiac hypertrophy is usually characterized morphologically by increases in the size of myofibrils and mitochondria, as well as by enlargement of mitochondria and nuclei. At this stage, while muscle cells are larger than normal, cellular organization is largely preserved. At a more advanced stage of cardiac hypertrophy, there are preferential increases in the size or number of specific organelles, such as mitochondria, and new contractile elements are added in localized areas of the cells, in an irregular manner. Cells subjected to long-standing hypertrophy show more obvious disruptions in cellular organization, including markedly enlarged nuclei with highly lobulated membranes, which displace adjacent myofibrils and cause breakdown of normal Z-band registration. The phrase "cardiac hypertrophy" is used to include all stages of the progression of this condition, characterized by various degrees of structural damage of the heart muscle, regardless of the underlying cardiac disorder. Hence, the term also includes physiological conditions instrumental in the development of cardiac hypertrophy, such as elevated blood pressure, aortic stenosis, or myocardial infarction.

"Heart failure" refers to an abnormality of cardiac function where the heart does not pump blood at the rate

needed for the requirements of metabolizing tissues. The heart failure can be caused by a number of factors, including ischemic, congenital, rheumatic, or idiopathic forms.

“Congestive heart failure” (CHF) is a progressive pathologic state where the heart is increasingly unable to supply adequate cardiac output (the volume of blood pumped by the heart over time) to deliver the oxygenated blood to peripheral tissues. As CHF progresses, structural and hemodynamic damages occur. While these damages have a variety of manifestations, one characteristic symptom is ventricular hypertrophy. CHF is a common end result of a number of various cardiac disorders.

“Myocardial infarction” generally results from atherosclerosis of the coronary arteries, often with superimposed coronary thrombosis. It may be divided into two major types: transmural infarcts, in which myocardial necrosis involves the full thickness of the ventricular wall, and subendocardial (nontransmural) infarcts, in which the necrosis involves the subendocardium, the intramural myocardium, or both, without extending all the way through the ventricular wall to the epicardium. Myocardial infarction is known to cause both a change in hemodynamic effects and an alteration in structure in the damaged and healthy zones of the heart. Thus, for example, myocardial infarction reduces the maximum cardiac output and the stroke volume of the heart. Also associated with myocardial infarction is a stimulation of the DNA synthesis occurring in the interstice as well as an increase in the formation of collagen in the areas of the heart not affected.

As a result of the increased stress or strain placed on the heart in prolonged hypertension due, for example, to the increased total peripheral resistance, cardiac hypertrophy has long been associated with “hypertension”. A characteristic of the ventricle that becomes hypertrophic as a result of chronic pressure overload is an impaired diastolic performance. Fouad *et al.*, J. Am. Coll. Cardiol., 4: 1500-1506 (1984); Smith *et al.*, J. Am. Coll. Cardiol., 5: 869-874 (1985). A prolonged left ventricular relaxation has been detected in early essential hypertension, in spite of normal or supranormal systolic function. Hartford *et al.*, Hypertension, 6: 329-338 (1984). However, there is no close parallelism between blood pressure levels and cardiac hypertrophy. Although improvement in left ventricular function in response to antihypertensive therapy has been reported in humans, patients variously treated with a diuretic (hydrochlorothiazide), a  $\beta$ -blocker (propranolol), or a calcium channel blocker (diltiazem), have shown reversal of left ventricular hypertrophy, without improvement in diastolic function. Inouye *et al.*, Am. J. Cardiol., 53: 1583-7 (1984).

Another complex cardiac disease associated with cardiac hypertrophy is “hypertrophic cardiomyopathy”. This condition is characterized by a great diversity of morphologic, functional, and clinical features (Maron *et al.*, N. Engl. J. Med., 316: 780-789 (1987); Spirito *et al.*, N. Engl. J. Med., 320: 749-755 (1989); Louie and Edwards, Prog. Cardiovasc. Dis., 36: 275-308 (1994); Wagle *et al.*, Circulation, 92: 1680-1692 (1995)), the heterogeneity of which is accentuated by the fact that it afflicts patients of all ages. Spirito *et al.*, N. Engl. J. Med., 336: 775-785 (1997). The causative factors of hypertrophic cardiomyopathy are also diverse and little understood. In general, mutations in genes encoding sarcomeric proteins are associated with hypertrophic cardiomyopathy. Recent data suggest that  $\beta$ -myosin heavy chain mutations may account for approximately 30 to 40 percent of cases of familial hypertrophic cardiomyopathy. Watkins *et al.*, N. Engl. J. Med., 326: 1108-1114 (1992); Schwartz *et al.*, Circulation, 91: 532-540 (1995); Marian and Roberts, Circulation, 92: 1336-1347 (1995); Thierfelder *et al.*, Cell, 77: 701-712

(1994); Watkins *et al.*, Nat. Gen., **11**: 434-437 (1995). Besides  $\beta$ -myosin heavy chain, other locations of genetic mutations include cardiac troponin T, alpha topomyosin, cardiac myosin binding protein C, essential myosin light chain, and regulatory myosin light chain. See, Malik and Watkins, Curr. Opin. Cardiol., **12**: 295-302 (1997).

5       Supravalvular "aortic stenosis" is an inherited vascular disorder characterized by narrowing of the ascending aorta, but other arteries, including the pulmonary arteries, may also be affected. Untreated aortic stenosis may lead to increased intracardiac pressure resulting in myocardial hypertrophy and eventually heart failure and death. The pathogenesis of this disorder is not fully understood, but hypertrophy and possibly hyperplasia of medial smooth muscle are prominent features of this disorder. It has been reported that molecular variants of the elastin gene are involved in the development and pathogenesis of aortic stenosis. U.S. Patent No. 5,650,282 issued July 10       22, 1997.

      "Valvular regurgitation" occurs as a result of heart diseases resulting in disorders of the cardiac valves. Various diseases, like rheumatic fever, can cause the shrinking or pulling apart of the valve orifice, while other diseases may result in endocarditis, an inflammation of the endocardium or lining membrane of the atrioventricular orifices and operation of the heart. Defects such as the narrowing of the valve stenosis or the defective closing of 15       the valve result in an accumulation of blood in the heart cavity or regurgitation of blood past the valve. If uncorrected, prolonged valvular stenosis or insufficiency may result in cardiac hypertrophy and associated damage to the heart muscle, which may eventually necessitate valve replacement.

      The treatment of all these, and other cardiovascular, endothelial and angiogenic disorders, which may or may not be accompanied by cardiac hypertrophy, is encompassed by the present invention.

20       The terms "cancer", "cancerous", and "malignant" refer to or describe the physiological condition in mammals that is typically characterized by unregulated cell growth. Examples of cancer include but are not limited to, carcinoma including adenocarcinoma, lymphoma, blastoma, melanoma, sarcoma, and leukemia. More particular examples of such cancers include squamous cell cancer, small-cell lung cancer, non-small cell lung cancer, gastrointestinal cancer, Hodgkin's and non-Hodgkin's lymphoma, pancreatic cancer, glioblastoma, cervical cancer, 25       ovarian cancer, liver cancer such as hepatic carcinoma and hepatoma, bladder cancer, breast cancer, colon cancer, colorectal cancer, endometrial carcinoma, salivary gland carcinoma, kidney cancer such as renal cell carcinoma and Wilms' tumors, basal cell carcinoma, melanoma, prostate cancer, vulval cancer, thyroid cancer, testicular cancer, esophageal cancer, and various types of head and neck cancer. The preferred cancers for treatment herein are breast, colon, lung, melanoma, ovarian, and others involving vascular tumors as noted above.

30       The term "cytotoxic agent" as used herein refers to a substance that inhibits or prevents the function of cells and/or causes destruction of cells. The term is intended to include radioactive isotopes (*e.g.*,  $^{131}\text{I}$ ,  $^{125}\text{I}$ ,  $^{90}\text{Y}$ , and  $^{186}\text{Re}$ ), chemotherapeutic agents, and toxins such as enzymatically active toxins of bacterial, fungal, plant, or animal origin, or fragments thereof.

35       A "chemotherapeutic agent" is a chemical compound useful in the treatment of cancer. Examples of chemotherapeutic agents include alkylating agents, folic acid antagonists, anti-metabolites of nucleic acid metabolism, antibiotics, pyrimidine analogs, 5-fluorouracil, cisplatin, purine nucleosides, amines, amino acids, triazol nucleosides, or corticosteroids. Specific examples include Adriamycin, Doxorubicin, 5-Fluorouracil,

Cytosine arabinoside ("Ara-C"), Cyclophosphamide, Thiotepa, Busulfan, Cytosin, Taxol, Toxotere, Methotrexate, Cisplatin, Melphalan, Vinblastine, Bleomycin, Etoposide, Ifosfamide, Mitomycin C, Mitoxantrone, Vincristine, Vinorelbine, Carboplatin, Teniposide, Daunomycin, Carminomycin, Aminopterin, Dactinomycin, Mitomycins, Esperamicins (see U.S. Pat. No. 4,675,187), Melphalan, and other related nitrogen mustards. Also included in this definition are hormonal agents that act to regulate or inhibit hormone action on tumors, such as tamoxifen and onapristone.

A "growth-inhibitory agent" when used herein refers to a compound or composition that inhibits growth of a cell, such as an Wnt-overexpressing cancer cell, either *in vitro* or *in vivo*. Thus, the growth-inhibitory agent is one which significantly reduces the percentage of malignant cells in S phase. Examples of growth-inhibitory agents include agents that block cell cycle progression (at a place other than S phase), such as agents that induce G1 arrest and M-phase arrest. Classical M-phase blockers include the vincas (vincristine and vinblastine), taxol, and topo II inhibitors such as doxorubicin, daunorubicin, etoposide, and bleomycin. Those agents that arrest G1 also spill over into S-phase arrest, for example, DNA alkylating agents such as tamoxifen, prednisone, dacarbazine, mechlorethamine, cisplatin, methotrexate, 5-fluorouracil, and ara-C. Further information can be found in The Molecular Basis of Cancer, Mendelsohn and Israel, eds., Chapter 1, entitled "Cell cycle regulation, oncogenes, and antineoplastic drugs" by Murakami *et al.* (WB Saunders: Philadelphia, 1995), especially p. 13. Additional examples include tumor necrosis factor (TNF), an antibody capable of inhibiting or neutralizing the angiogenic activity of acidic or basic FGF or hepatocyte growth factor (HGF), an antibody capable of inhibiting or neutralizing the coagulant activities of tissue factor, protein C, or protein S (see, WO 91/01753, published 21 February 1991), or an antibody capable of binding to HER2 receptor (WO 89/06692), such as the 4D5 antibody (and functional equivalents thereof) (e.g., WO 92/22653).

"Treatment" is an intervention performed with the intention of preventing the development or altering the pathology of a cardiovascular, endothelial, and angiogenic disorder. The concept of treatment is used in the broadest sense, and specifically includes the prevention (prophylaxis), moderation, reduction, and curing of cardiovascular, endothelial, and angiogenic disorders of any stage. Accordingly, "treatment" refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) or ameliorate a cardiovascular, endothelial, and angiogenic disorder such as hypertrophy. Those in need of treatment include those already with the disorder as well as those prone to have the disorder or those in whom the disorder is to be prevented. The disorder may result from any cause, including idiopathic, cardiopathic, or myopathic causes, or ischemia or ischemic insults, such as myocardial infarction.

"Chronic" administration refers to administration of the agent(s) in a continuous mode as opposed to an acute mode, so as to maintain the initial effect, such as an anti-hypertrophic effect, for an extended period of time.

"Mammal" for purposes of treatment refers to any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, horses, cats, cows, sheep, pigs, etc. Preferably, the mammal is human.

Administration "in combination with" one or more further therapeutic agents includes simultaneous (concurrent) and consecutive administration in any order.

The phrase "cardiovascular, endothelial or angiogenic agents" refers generically to any drug that acts in treating cardiovascular, endothelial, and angiogenic disorders. Examples of cardiovascular agents are those that promote vascular homeostasis by modulating blood pressure, heart rate, heart contractility, and endothelial and smooth muscle biology, all of which factors have a role in cardiovascular disease. Specific examples of these include angiotensin-II receptor antagonists; endothelin receptor antagonists such as, for example, BOSENTAN<sup>TM</sup> and MOXONODIN<sup>TM</sup>; interferon-gamma (IFN- $\gamma$ ); des-aspartate-angiotensin I; thrombolytic agents, *e.g.*, streptokinase, urokinase, t-PA, and a t-PA variant specifically designed to have longer half-life and very high fibrin specificity, TNK t-PA (a T103N, N117Q, KHRR(296-299)AAAA t-PA variant, Keyt *et al.*, Proc. Natl. Acad. Sci. USA, 91: 3670-3674 (1994)); inotropic or hypertensive agents such as digoxigenin and  $\beta$ -adrenergic receptor blocking agents, *e.g.*, propranolol, timolol, tertalolol, carteolol, nadolol, betaxolol, penbutolol, acetobutolol, atenolol, metoprolol, and carvedilol; angiotensin converting enzyme (ACE) inhibitors, *e.g.*, quinapril, captopril, enalapril, ramipril, benazepril, fosinopril, and lisinopril; diuretics, *e.g.*, chlorothiazide, hydrochlorothiazide, hydroflumethazide, methylchlorothiazide, benzthiazide, dichlorphenamide, acetazolamide, and indapamide; and calcium channel blockers, *e.g.*, diltiazem, nifedipine, verapamil, nicardipine. One preferred category of this type is a therapeutic agent used for the treatment of cardiac hypertrophy or of a physiological condition instrumental in the development of cardiac hypertrophy, such as elevated blood pressure, aortic stenosis, or myocardial infarction.

"Angiogenic agents" and "endothelial agents" are active agents that promote angiogenesis and/or endothelial cell growth, or, if applicable, vasculogenesis. This would include factors that accelerate wound healing, such as growth hormone, insulin-like growth factor-I (IGF-I), VEGF, VIGF, PDGF, epidermal growth factor (EGF), CTGF and members of its family, FGF, and TGF- $\alpha$  and TGF- $\beta$ .

"Angiostatic agents" are active agents that inhibit angiogenesis or vasculogenesis or otherwise inhibit or prevent growth of cancer cells. Examples include antibodies or other antagonists to angiogenic agents as defined above, such as antibodies to VEGF. They additionally include cytotherapeutic agents such as cytotoxic agents, chemotherapeutic agents, growth-inhibitory agents, apoptotic agents, and other agents to treat cancer, such as anti-HER-2, anti-CD20, and other bioactive and organic chemical agents.

In a pharmacological sense, in the context of the present invention, a "therapeutically effective amount" of an active agent such as a PRO polypeptide or agonist or antagonist thereto or an anti-PRO antibody, refers to an amount effective in the treatment of a cardiovascular, endothelial or angiogenic disorder in a mammal and can be determined empirically.

As used herein, an "effective amount" of an active agent such as a PRO polypeptide or agonist or antagonist thereto or an anti-PRO antibody, refers to an amount effective for carrying out a stated purpose, wherein such amounts may be determined empirically for the desired effect.

The terms "PRO polypeptide" and "PRO" as used herein and when immediately followed by a numerical designation refer to various polypeptides, wherein the complete designation (*i.e.*, PRO/number) refers to specific polypeptide sequences as described herein. The terms "PRO/number polypeptide" and "PRO/number" wherein the term "number" is provided as an actual numerical designation as used herein encompass native sequence polypeptides and polypeptide variants (which are further defined herein). The PRO polypeptides described herein

may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods.

5 A "native sequence PRO polypeptide" comprises a polypeptide having the same amino acid sequence as the corresponding PRO polypeptide derived from nature. Such native sequence PRO polypeptides can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence PRO polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of the specific PRO polypeptide (*e.g.*, an extracellular domain sequence), naturally-occurring variant forms (*e.g.*, alternatively spliced forms) and naturally-occurring allelic variants of the polypeptide. In various embodiments of the invention, the native sequence PRO polypeptides disclosed herein are mature or full-length native sequence polypeptides comprising the full-length amino acids sequences shown in the accompanying figures. Start and stop codons are shown in bold font and underlined in the figures. However, while the PRO polypeptide disclosed in the accompanying figures are shown to begin with methionine residues designated herein as amino acid position 1 in the figures, it is conceivable and possible that other methionine residues located either upstream or downstream from the amino acid position 1 in the figures may be employed as the starting amino acid residue for the PRO polypeptides.

15 The PRO polypeptide "extracellular domain" or "ECD" refers to a form of the PRO polypeptide which is essentially free of the transmembrane and cytoplasmic domains. Ordinarily, a PRO polypeptide ECD will have less than 1% of such transmembrane and/or cytoplasmic domains and preferably, will have less than 0.5% of such domains. It will be understood that any transmembrane domains identified for the PRO polypeptides of the present invention are identified pursuant to criteria routinely employed in the art for identifying that type of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified herein. Optionally, therefore, an extracellular domain of a PRO polypeptide may contain from about 5 or fewer amino acids on either side of the transmembrane domain/extracellular domain boundary as identified in the Examples or specification and such polypeptides, with or without the associated signal peptide, and nucleic acid encoding them, are contemplated by the present invention.

25 The approximate location of the "signal peptides" of the various PRO polypeptides disclosed herein are shown in the present specification and/or the accompanying figures. It is noted, however, that the C-terminal boundary of a signal peptide may vary, but most likely by no more than about 5 amino acids on either side of the signal peptide C-terminal boundary as initially identified herein, wherein the C-terminal boundary of the signal peptide may be identified pursuant to criteria routinely employed in the art for identifying that type of amino acid sequence element (*e.g.*, Nielsen *et al.*, Prot. Eng., 10:1-6 (1997) and von Heinje *et al.*, Nucl. Acids Res., 14:4683-4690 (1986)). Moreover, it is also recognized that, in some cases, cleavage of a signal sequence from a secreted polypeptide is not entirely uniform, resulting in more than one secreted species. These mature polypeptides, where the signal peptide is cleaved within no more than about 5 amino acids on either side of the C-terminal boundary of the signal peptide as identified herein, and the polynucleotides encoding them, are contemplated by the present invention.

35 "PRO polypeptide variant" means an active PRO polypeptide as defined above or below having at least about 80% amino acid sequence identity with a full-length native sequence PRO polypeptide sequence as disclosed



herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Such PRO polypeptide variants include, for instance, PRO polypeptides wherein one or more amino acid residues are added, or deleted, at the N- or C-terminus of the full-length native amino acid sequence. Ordinarily, a PRO polypeptide variant will have at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97% or 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, PRO variant polypeptides are at least about 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150 or 200 amino acids in length and alternatively at least about 300 amino acids in length, or more.

"Percent (%) amino acid sequence identity" with respect to the PRO polypeptide sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in a PRO sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN, ALIGN-2 or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full-length of the sequences being compared. For purposes herein, however, % amino acid sequence identity values are obtained as described below by using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc., and the source code shown in Table 1 has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

For purposes herein, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

$$100 \text{ times the fraction } X/Y$$

where X is the number of amino acid residues scored as identical matches by the sequence alignment program

ALIGN-2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A. As examples of % amino acid sequence identity calculations, Tables 2-3 demonstrate how to calculate the % amino acid sequence identity of the amino acid sequence designated "Comparison Protein" to the amino acid sequence designated "PRO".

Unless specifically stated otherwise, all % amino acid sequence identity values used herein are obtained as described above using the ALIGN-2 sequence comparison computer program. However, % amino acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul *et al.*, Nucleic Acids Res., 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov>, or otherwise obtained from the National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

$$100 \text{ times the fraction } X/Y$$

where X is the number of amino acid residues scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A.

In addition, % amino acid sequence identity may also be determined using the WU-BLAST-2 computer program (Altschul *et al.*, Methods in Enzymology, 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, *i.e.*, the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. For purposes herein, a % amino acid sequence identity value is determined by dividing (a) the number of matching identical amino acids residues between the amino acid sequence of the PRO polypeptide of interest having a sequence derived from the native PRO polypeptide and the comparison amino acid sequence of interest (*i.e.*, the sequence against which the PRO polypeptide of interest is being compared which may be a PRO variant polypeptide) as determined by WU-BLAST-2 by (b) the total number of amino acid residues of the PRO polypeptide of interest. For example, in the statement "a polypeptide comprising an amino acid sequence A which

has or having at least 80% amino acid sequence identity to the amino acid sequence B", the amino acid sequence A is the comparison amino acid sequence of interest and the amino acid sequence B is the amino acid sequence of the PRO polypeptide of interest.

5 "PRO variant polynucleotide" or "PRO variant nucleic acid sequence" means a nucleic acid molecule which encodes an active PRO polypeptide as defined below and which has at least about 80% nucleic acid sequence identity with a nucleotide acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, a PRO variant  
10 polynucleotide will have at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97% or 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity with a nucleic acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal sequence, as disclosed herein or  
15 any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Variants do not encompass the native nucleotide sequence.

Ordinarily, PRO variant polynucleotides are at least about 30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 450, or 600 nucleotides in length and alternatively at least about 900 nucleotides in length, or more.

"Percent (%) nucleic acid sequence identity" with respect to the PRO polypeptide-encoding nucleic acid  
20 sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in a PRO polypeptide-encoding nucleic acid sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN, ALIGN-2 or Megalign  
25 (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full-length of the sequences being compared. For purposes herein, however, % nucleic acid sequence identity values are obtained as described below by using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1. The ALIGN-2 sequence comparison computer program was authored by  
30 Genentech, Inc., and the source code shown in Table 1 has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2  
35 program and do not vary.

For purposes herein, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that

has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

$$100 \text{ times the fraction } W/Z$$

5

where W is the number of nucleotides scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C. As examples of %  
10 nucleic acid sequence identity calculations, Tables 4-5 demonstrate how to calculate the % nucleic acid sequence identity of the nucleic acid sequence designated "Comparison DNA" to the nucleic acid sequence designated "PRO-DNA".

Unless specifically stated otherwise, all % nucleic acid sequence identity values used herein are obtained as described above using the ALIGN-2 sequence comparison computer program. However, % nucleic acid  
15 sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul *et al.*, Nucleic Acids Res., 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov> or otherwise obtained from the National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low  
20 complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence  
25 identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

$$100 \text{ times the fraction } W/Z$$

where W is the number of nucleotides scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C.

In addition, % nucleic acid sequence identity values may also be generated using the WU-BLAST-2 computer program (Altschul *et al.*, Methods in Enzymology, 266:460-480 (1996)). Most of the WU-BLAST-2  
35 search parameters are set to the default values. Those not set to default values, *i.e.*, the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. For purposes herein, a % nucleic acid sequence identity value is determined by dividing (a)

the number of matching identical nucleotides between the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest having a sequence derived from the native sequence PRO polypeptide-encoding nucleic acid and the comparison nucleic acid molecule of interest (i.e., the sequence against which the PRO polypeptide-encoding nucleic acid molecule of interest is being compared which may be a variant PRO polynucleotide) as determined by WU-BLAST-2 by (b) the total number of nucleotides of the PRO polypeptide-encoding nucleic acid molecule of interest. For example, in the statement "an isolated nucleic acid molecule comprising a nucleic acid sequence A which has or having at least 80% nucleic acid sequence identity to the nucleic acid sequence B", the nucleic acid sequence A is the comparison nucleic acid molecule of interest and the nucleic acid sequence B is the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest.

In other embodiments, PRO variant polynucleotides are nucleic acid molecules that encode an active PRO polypeptide and which are capable of hybridizing, preferably under stringent hybridization and wash conditions, to nucleotide sequences encoding the full-length PRO polypeptide as shown in the specification herein and accompanying figures. PRO variant polypeptides may be those that are encoded by a PRO variant polynucleotide.

"Isolated", when used to describe the various polypeptides disclosed herein, means a polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Preferably, the isolated polypeptide is free of association with all components with which it is naturally associated. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the PRO natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

An "isolated" nucleic acid molecule encoding a PRO polypeptide or an "isolated" nucleic acid molecule encoding an anti-PRO antibody is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the PRO-encoding nucleic acid or the natural source of the anti-PRO-encoding nucleic acid. Preferably, the isolated nucleic acid is free of association with all components with which it is naturally associated. An isolated PRO-encoding nucleic acid molecule or an isolated anti-PRO-encoding nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated nucleic acid molecules therefore are distinguished from the PRO-encoding nucleic acid molecule or from the anti-PRO-encoding nucleic acid molecule as it exists in natural cells. However, an isolated nucleic acid molecule encoding a PRO polypeptide or an isolated nucleic acid molecule encoding an anti-PRO antibody includes PRO-nucleic acid molecules or anti-PRO-nucleic acid molecules contained in cells that ordinarily express PRO polypeptides or anti-PRO antibodies where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked

coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize, for example, promoters, polyadenylation signals, and enhancers.

5 Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a PRO polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, 10 contiguous and in the same reading frame. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

"Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. 15 In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature that can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower 20 temperatures less so. For additional details and explanation of stringency of hybridization reactions, *see*, Ausubel *et al.*, Current Protocols in Molecular Biology (Wiley Interscience Publishers, 1995).

"Stringent conditions" or "high-stringency conditions", as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example, 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as 25 formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1 x SSC containing EDTA 30 at 55°C.

"Moderately-stringent conditions" may be identified as described by Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Press, 1989), and include the use of washing solution and hybridization conditions (*e.g.*, temperature, ionic strength, and % SDS) less stringent than those described above. 35 An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters

in 1 x SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

The modifier "epitope-tagged" when used herein refers to a chimeric polypeptide comprising a PRO polypeptide fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino acid residues).

"Active" or "activity" in the context of PRO variants refers to form(s) of PRO proteins that retain the biologic and/or immunologic activities of a native or naturally-occurring PRO polypeptide.

"Biological activity" in the context of a molecule that antagonizes a PRO polypeptide that can be identified by the screening assays disclosed herein (*e.g.*, an organic or inorganic small molecule, peptide, etc.) is used to refer to the ability of such molecules to bind or complex with the PRO polypeptide identified herein, or otherwise interfere with the interaction of the PRO polypeptide with other cellular proteins or otherwise inhibits the transcription or translation of the PRO polypeptide. Particularly preferred biological activity includes cardiac hypertrophy, activity that acts on systemic disorders that affect vessels, such as diabetes mellitus, as well as diseases of the arteries, capillaries, veins, and/or lymphatics, and cancer.

The term "antagonist" is used in the broadest sense, and includes any molecule that partially or fully blocks, inhibits, or neutralizes one or more of the biological activities of a native PRO polypeptide disclosed herein, for example, if applicable, its mitogenic or angiogenic activity. Antagonists of a PRO polypeptide may act by interfering with the binding of a PRO polypeptide to a cellular receptor, by incapacitating or killing cells that have been activated by a PRO polypeptide, or by interfering with vascular endothelial cell activation after binding of a PRO polypeptide to a cellular receptor. All such points of intervention by a PRO polypeptide antagonist shall be considered equivalent for purposes of this invention. The antagonists inhibit the mitogenic, angiogenic, or other biological activity of PRO polypeptides, and thus are useful for the treatment of diseases or disorders characterized by undesirable excessive neovascularization, including by way of example tumors, and especially solid malignant tumors, rheumatoid arthritis, psoriasis, atherosclerosis, diabetic and other retinopathies, retrolental fibroplasia, age-related macular degeneration, neovascular glaucoma, hemangiomas, thyroid hyperplasias (including Grave's disease), corneal and other tissue transplantation, and chronic inflammation. The antagonists also are useful for the treatment of diseases or disorders characterized by undesirable excessive vascular permeability, such as edema associated with brain tumors, ascites associated with malignancies, Meigs' syndrome, lung inflammation, nephrotic syndrome, pericardial effusion (such as that associated with pericarditis), and pleural effusion. In a similar manner, the term "agonist" is used in the broadest sense and includes any molecule that mimics a biological activity of a native PRO polypeptide disclosed herein. Suitable agonist or antagonist molecules specifically include agonist or antagonist antibodies or antibody fragments, fragments, or amino acid sequence variants of native PRO polypeptides, peptides, small organic molecules, etc.

A "small molecule" is defined herein to have a molecular weight below about 500 daltons.

The term "PRO polypeptide receptor" as used herein refers to a cellular receptor for a PRO polypeptide, ordinarily a cell-surface receptor found on vascular endothelial cells, as well as variants thereof that retain the ability to bind a PRO polypeptide.

"Antibodies" (Abs) and "immunoglobulins" (Igs) are glycoproteins having the same structural characteristics. While antibodies exhibit binding specificity to a specific antigen, immunoglobulins include both antibodies and other antibody-like molecules that lack antigen specificity. Polypeptides of the latter kind are, for example, produced at low levels by the lymph system and at increased levels by myelomas. The term "antibody" is used in the broadest sense and specifically covers, without limitation, intact monoclonal antibodies, polyclonal antibodies, multispecific antibodies (*e.g.*, bispecific antibodies) formed from at least two intact antibodies, and antibody fragments, so long as they exhibit the desired biological activity.

"Native antibodies" and "native immunoglobulins" are usually heterotetrameric glycoproteins of about 150,000 daltons, composed of two identical light (L) chains and two identical heavy (H) chains. Each light chain is linked to a heavy chain by one covalent disulfide bond, while the number of disulfide linkages varies among the heavy chains of different immunoglobulin isotypes. Each heavy and light chain also has regularly spaced intrachain disulfide bridges. Each heavy chain has at one end a variable domain ( $V_H$ ) followed by a number of constant domains. Each light chain has a variable domain at one end ( $V_L$ ) and a constant domain at its other end; the constant domain of the light chain is aligned with the first constant domain of the heavy chain, and the light-chain variable domain is aligned with the variable domain of the heavy chain. Particular amino acid residues are believed to form an interface between the light- and heavy-chain variable domains.

The term "variable" refers to the fact that certain portions of the variable domains differ extensively in sequence among antibodies and are used in the binding and specificity of each particular antibody to and for its particular antigen. However, the variability is not evenly distributed throughout the variable domains of antibodies. It is concentrated in three segments called complementarity-determining regions (CDRs) or hypervariable regions both in the light-chain and the heavy-chain variable domains. The more highly conserved portions of variable domains are called the framework regions (FR). The variable domains of native heavy and light chains each comprise four FR regions, largely adopting a  $\beta$ -sheet configuration, connected by three CDRs, which form loops connecting, and in some cases forming part of, the  $\beta$ -sheet structure. The CDRs in each chain are held together in close proximity by the FR regions and, with the CDRs from the other chain, contribute to the formation of the antigen-binding site of antibodies. See, Kabat *et al.*, NIH Publ. No.91-3242, Vol. I, pages 647-669 (1991). The constant domains are not involved directly in binding an antibody to an antigen, but exhibit various effector functions, such as participation of the antibody in antibody-dependent cellular toxicity.

"Antibody fragments" comprise a portion of an intact antibody, preferably the antigen-binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', F(ab')<sub>2</sub>, and Fv fragments; diabodies; linear antibodies (Zapata *et al.*, Protein Eng., **8**(10): 1057-1062 (1995)); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, each with a single antigen-binding site, and a residual "Fc" fragment, whose name reflects its ability to crystallize



readily. Pepsin treatment yields an  $F(ab')_2$  fragment that has two antigen-combining sites and is still capable of cross-linking antigen.

"Fv" is the minimum antibody fragment that contains a complete antigen-recognition and -binding site. This region consists of a dimer of one heavy- and one light-chain variable domain in tight, non-covalent association. It is in this configuration that the three CDRs of each variable domain interact to define an antigen-binding site on the surface of the  $V_H$ - $V_L$  dimer. Collectively, the six CDRs confer antigen-binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

The Fab fragment also contains the constant domain of the light chain and the first constant domain (CH1) of the heavy chain. Fab' fragments differ from Fab fragments by the addition of a few residues at the carboxy terminus of the heavy chain CH1 domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group.  $F(ab')_2$  antibody fragments originally were produced as pairs of Fab' fragments that have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

The "light chains" of antibodies (immunoglobulins) from any vertebrate species can be assigned to one of two clearly distinct types, called kappa ( $\kappa$ ) and lambda ( $\lambda$ ), based on the amino acid sequences of their constant domains.

Depending on the amino acid sequence of the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are five major classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM; and several of these may be further divided into subclasses (isotypes), e.g., IgG1, IgG2, IgG3, IgG4, IgA, and IgA2. The heavy-chain constant domains that correspond to the different classes of immunoglobulins are called  $\alpha$ ,  $\delta$ ,  $\epsilon$ ,  $\gamma$ , and  $\mu$ , respectively. The subunit structures and three-dimensional configurations of different classes of immunoglobulins are well known.

The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site. Furthermore, in contrast to conventional (polyclonal) antibody preparations that typically include different antibodies directed against different determinants (epitopes), each monoclonal antibody is directed against a single determinant on the antigen. In addition to their specificity, the monoclonal antibodies are advantageous in that they are synthesized by the hybridoma culture, uncontaminated by other immunoglobulins. The modifier "monoclonal" indicates the character of the antibody as being obtained from a substantially homogeneous population of antibodies, and is not to be construed as requiring production of the antibody by any particular method. For example, the monoclonal antibodies to be used in accordance with the present invention may be made by the hybridoma method first described by Kohler *et al.*, Nature, 256: 495 (1975), or may be made by recombinant DNA methods (see, e.g., U.S. Patent No. 4,816,567). The "monoclonal antibodies" may also be isolated from phage antibody libraries using the techniques described in Clackson *et al.*, Nature, 352: 624-628 (1991) and Marks *et al.*, J. Mol. Biol., 222: 581-597 (1991), for example.

The monoclonal antibodies herein specifically include "chimeric" antibodies (immunoglobulins) in which a portion of the heavy and/or light chain is identical with or homologous to corresponding sequences in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is identical with or homologous to corresponding sequences in antibodies derived from another species or belonging to another antibody class or subclass, as well as fragments of such antibodies, so long as they exhibit the desired biological activity. U.S. Patent No. 4,816,567; Morrison *et al.*, Proc. Natl. Acad. Sci. USA, **81**: 6851-6855 (1984).

"Humanized" forms of non-human (*e.g.*, murine) antibodies are chimeric immunoglobulins, immunoglobulin chains, or fragments thereof (such as Fv, Fab, Fab', F(ab')<sub>2</sub>, or other antigen-binding subsequences of antibodies) that contain minimal sequence derived from non-human immunoglobulin. For the most part, humanized antibodies are human immunoglobulins (recipient antibody) in which residues from a CDR of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity, and capacity. In some instances, Fv FR residues of the human immunoglobulin are replaced by corresponding non-human residues. Furthermore, humanized antibodies may comprise residues that are found neither in the recipient antibody nor in the imported CDR or framework sequences. These modifications are made to further refine and maximize antibody performance. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin sequence. The humanized antibody preferably also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin. For further details, see Jones *et al.*, Nature, **321**: 522-525 (1986); Reichmann *et al.*, Nature, **332**: 323-329 (1988); and Presta, Curr. Op. Struct. Biol., **2**: 593-596 (1992). The humanized antibody includes a PRIMATIZED™ antibody wherein the antigen-binding region of the antibody is derived from an antibody produced by immunizing macaque monkeys with the antigen of interest.

"Single-chain Fv" or "sFv" antibody fragments comprise the V<sub>H</sub> and V<sub>L</sub> domains of an antibody, wherein these domains are present in a single polypeptide chain. Preferably, the Fv polypeptide further comprises a polypeptide linker between the V<sub>H</sub> and V<sub>L</sub> domains that enables the sFv to form the desired structure for antigen binding. For a review of sFv see, Pluckthun in The Pharmacology of Monoclonal Antibodies, Vol. 113, Rosenberg and Moore, eds. (Springer-Verlag: New York, 1994), pp. 269-315.

The term "diabodies" refers to small antibody fragments with two antigen-binding sites, which fragments comprise a heavy-chain variable domain (V<sub>H</sub>) connected to a light-chain variable domain (V<sub>L</sub>) in the same polypeptide chain (V<sub>H</sub> - V<sub>L</sub>). By using a linker that is too short to allow pairing between the two domains on the same chain, the domains are forced to pair with the complementary domains of another chain and create two antigen-binding sites. Diabodies are described more fully in, for example, EP 404,097; WO 93/11161; and Hollinger *et al.*, Proc. Natl. Acad. Sci. USA, **90**: 6444-6448 (1993).

An "isolated" antibody is one that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would interfere

with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody will be purified (1) to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under reducing or nonreducing conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody *in situ* within recombinant cells, since at least one component of the antibody's natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

An antibody that "specifically binds to" or is "specific for" a particular polypeptide or an epitope on a particular polypeptide is one that binds to that particular polypeptide or epitope on a particular polypeptide without substantially binding to any other polypeptide or polypeptide epitope.

The word "label" when used herein refers to a detectable compound or other composition that is conjugated directly or indirectly to the antibody so as to generate a "labeled" antibody. The label may be detectable by itself (e.g., radioisotope labels or fluorescent labels) or, in the case of an enzymatic label, may catalyze chemical alteration of a substrate compound or composition that is detectable. Radionuclides that can serve as detectable labels include, for example, I-131, I-123, I-125, Y-90, Re-188, At-211, Cu-67, Bi-212, and Pd-109. The label may also be a non-detectable entity such as a toxin.

By "solid phase" is meant a non-aqueous matrix to which an antibody of the present invention can adhere. Examples of solid phases encompassed herein include those formed partially or entirely of glass (e.g., controlled pore glass), polysaccharides (e.g., agarose), polyacrylamides, polystyrene, polyvinyl alcohol and silicones. In certain embodiments, depending on the context, the solid phase can comprise the well of an assay plate; in others it is a purification column (e.g., an affinity chromatography column). This term also includes a discontinuous solid phase of discrete particles, such as those described in U.S. Patent No. 4,275,149.

A "liposome" is a small vesicle composed of various types of lipids, phospholipids and/or surfactant that is useful for delivery of a drug (such as the PRO polypeptide or antibodies thereto disclosed herein) to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes.

As used herein, the term "immunoadhesin" designates antibody-like molecules that combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity that is other than the antigen recognition and binding site of an antibody (i.e., is "heterologous"), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any immunoglobulin, such as IgG-1, IgG-2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2), IgE, IgD, or IgM.

As shown below, Table 1 provides the complete source code for the ALIGN-2 sequence comparison computer program. This source code may be routinely compiled for use on a UNIX operating system to provide

the ALIGN-2 sequence comparison computer program.

5 In addition, Tables 2-5 show hypothetical exemplifications for using the below described method to determine % amino acid sequence identity (Tables 2-3) and % nucleic acid sequence identity (Tables 4-5) using the ALIGN-2 sequence comparison computer program, wherein "PRO" represents the amino acid sequence of a  
10 hypothetical PRO polypeptide of interest, "Comparison Protein" represents the amino acid sequence of a polypeptide against which the "PRO" polypeptide of interest is being compared, "PRO-DNA" represents a hypothetical PRO-encoding nucleic acid sequence of interest, "Comparison DNA" represents the nucleotide sequence of a nucleic acid molecule against which the "PRO-DNA" nucleic acid molecule of interest is being compared, "X", "Y", and "Z" each represent different hypothetical amino acid residues and "N", "L" and "V" each represent different hypothetical nucleotides.

**Table 1**

```

/*
 *
 * C-C increased from 12 to 15
 * Z is average of EQ
 * B is average of ND
 * match with stop is _M; stop-stop = 0; J (joker) match = 0
 */
#define _M      -8      /* value of a match with a stop */

int _day[26][26] = {
/* A B C D E F G H I J K L M N O P Q R S T U V W X Y Z */
/* A */ { 2, 0, -2, 0, 0, 4, 1, -1, -1, 0, -1, -2, -1, 0, _M, 1, 0, -2, 1, 1, 0, 0, -6, 0, -3, 0},
/* B */ { 0, 3, -4, 3, 2, -5, 0, 1, -2, 0, 0, -3, -2, 2, _M, -1, 1, 0, 0, 0, 0, -2, -5, 0, -3, 1},
/* C */ { -2, -4, 15, -5, -5, -4, -3, -3, -2, 0, -5, -6, -5, -4, _M, -3, -5, -4, 0, -2, 0, -2, -8, 0, 0, -5},
/* D */ { 0, 3, -5, 4, 3, -6, 1, 1, -2, 0, 0, -4, -3, 2, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 2},
/* E */ { 0, 2, -5, 3, 4, -5, 0, 1, -2, 0, 0, -3, -2, 1, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 3},
/* F */ { -4, -5, -4, -6, -5, 9, -5, -2, 1, 0, -5, 2, 0, -4, _M, -5, -5, -4, -3, -3, 0, -1, 0, 0, 7, -5},
/* G */ { 1, 0, -3, 1, 0, -5, 5, -2, -3, 0, -2, -4, -3, 0, _M, -1, -1, -3, 1, 0, 0, -1, -7, 0, -5, 0},
/* H */ { -1, 1, -3, 1, 1, -2, -2, 6, -2, 0, 0, -2, -2, 2, _M, 0, 3, 2, -1, -1, 0, -2, -3, 0, 0, 2},
/* I */ { -1, -2, -2, -2, -2, 1, -3, -2, 5, 0, -2, 2, 2, -2, _M, -2, -2, -2, -1, 0, 0, 4, -5, 0, -1, -2},
/* J */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */ { -1, 0, -5, 0, 0, -5, -2, 0, -2, 0, 5, -3, 0, 1, _M, -1, 1, 3, 0, 0, 0, -2, -3, 0, -4, 0},
/* L */ { -2, -3, -6, -4, -3, 2, -4, -2, 2, 0, -3, 6, 4, -3, _M, -3, -2, -3, -3, -1, 0, 2, -2, 0, -1, -2},
/* M */ { -1, -2, -5, -3, -2, 0, -3, -2, 2, 0, 0, 4, 6, -2, _M, -2, -1, 0, -2, -1, 0, 2, -4, 0, -2, -1},
/* N */ { 0, 2, -4, 2, 1, -4, 0, 2, -2, 0, 1, -3, -2, 2, _M, -1, 1, 0, 1, 0, 0, -2, -4, 0, -2, 1},
/* O */ { _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, 0, _M, _M, _M, _M, _M, _M, _M, _M, _M},
/* P */ { 1, -1, -3, -1, -1, -5, -1, 0, -2, 0, -1, -3, -2, -1, _M, 6, 0, 0, 1, 0, 0, -1, -6, 0, -5, 0},
/* Q */ { 0, 1, -5, 2, 2, -5, -1, 3, -2, 0, 1, -2, -1, 1, _M, 0, 4, 1, -1, -1, 0, -2, -5, 0, -4, 3},
/* R */ { -2, 0, -4, -1, -1, -4, -3, 2, -2, 0, 3, -3, 0, 0, _M, 0, 1, 6, 0, -1, 0, -2, 2, 0, -4, 0},
/* S */ { 1, 0, 0, 0, 0, -3, 1, -1, -1, 0, 0, -3, -2, 1, _M, 1, -1, 0, 2, 1, 0, -1, -2, 0, -3, 0},
/* T */ { 1, 0, -2, 0, 0, -3, 0, -1, 0, 0, 0, -1, -1, 0, _M, 0, -1, -1, 1, 3, 0, 0, -5, 0, -3, 0},
/* U */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */ { 0, -2, -2, -2, -2, -1, -1, -2, 4, 0, -2, 2, 2, -2, _M, -1, -2, -2, -1, 0, 0, 4, -6, 0, -2, -2},
/* W */ { -6, -5, -8, -7, -7, 0, -7, -3, -5, 0, -3, -2, -4, -4, _M, -6, -5, 2, -2, -5, 0, -6, 17, 0, 0, -6},
/* X */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */ { -3, -3, 0, -4, -4, 7, -5, 0, -1, 0, -4, -1, -2, -2, _M, -5, -4, -4, -3, -3, 0, -2, 0, 0, 10, -4},
/* Z */ { 0, 1, -5, 2, 3, -5, 0, 2, -2, 0, 0, -2, -1, 1, _M, 0, 3, 0, 0, 0, 0, -2, -6, 0, -4, 4}
};

```

**Table 1 (cont')**

```

/*
*/
#include <stdio.h>
#include <ctype.h>

#define MAXJMP      16      /* max jumps in a diag */
#define MAXGAP      24      /* don't continue to penalize gaps larger than this */
#define JMPS        1024    /* max jmps in an path */
#define MX          4       /* save if there's at least MX-1 bases since last jmp */

#define DMAT         3      /* value of matching bases */
#define DMIS         0      /* penalty for mismatched bases */
#define DINS0        8      /* penalty for a gap */
#define DINS1        1      /* penalty per base */
#define PINS0        8      /* penalty for a gap */
#define PINS1        4      /* penalty per residue */

struct jmp {
    short      n[MAXJMP];    /* size of jmp (neg for dely) */
    unsigned short x[MAXJMP]; /* base no. of jmp in seq x */
};
/* limits seq to 216 -1 */

struct diag {
    int      score;          /* score at last jmp */
    long     offset;         /* offset of prev block */
    short    jmp;           /* current jmp index */
    struct jmp jp;          /* list of jmps */
};

struct path {
    int      spc;            /* number of leading spaces */
    short    n[JMPS];        /* size of jmp (gap) */
    int      x[JMPS];        /* loc of jmp (last elem before gap) */
};

char      *ofile;           /* output file name */
char      *namex[2];        /* seq names: getseqs( ) */
char      *prog;            /* prog name for err msgs */
char      *seqx[2];         /* seqs: getseqs( ) */
int      dmax;              /* best diag: nw( ) */
int      dmax0;             /* final diag */
int      dna;               /* set if dna: main( ) */
int      endgaps;           /* set if penalizing end gaps */
int      gapx, gapy;        /* total gaps in seqs */
int      len0, len1;        /* seq lens */
int      ngapx, ngapy;      /* total size of gaps */
int      smax;              /* max score: nw( ) */
int      *xbm;              /* bitmap for matching */
long     offset;            /* current offset in jmp file */
struct    diag      *dx;    /* holds diagonals */
struct    path      pp[2];  /* holds path for seqs */

char      *calloc( ), *malloc( ), *index( ), *strcpy( );
char      *getseq( ), *g_calloc( );

```

Table 1 (cont')

```

/* Needleman-Wunsch alignment program
*
* usage: progs file1 file2
* where file1 and file2 are two dna or two protein sequences.
* The sequences can be in upper- or lower-case and may contain ambiguity
* Any lines beginning with ';', '>' or '<' are ignored
* Max file length is 65535 (limited by unsigned short x in the jmp struct)
* A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
* Output is in the file "align.out"
*
* The program may create a tmp file in /tmp to hold info about traceback.
* Original version developed under BSD 4.3 on a vax 8650
*/
#include "nw.h"
#include "day.h"

static  _dbval[26] = {
1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};

static  _pbval[26] = {
1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

main(ac, av)                                main
{
    int      ac;
    char     *av[];

    prog = av[0];
    if (ac != 3) {
        fprintf(stderr, "usage: %s file1 file2\n", prog);
        fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
        fprintf(stderr, "The sequences can be in upper- or lower-case\n");
        fprintf(stderr, "Any lines beginning with ';', '>' or '<' are ignored\n");
        fprintf(stderr, "Output is in the file \"align.out\"\n");
        exit(1);
    }
    namex[0] = av[1];
    namex[1] = av[2];
    seqx[0] = getseq(namex[0], &len0);
    seqx[1] = getseq(namex[1], &len1);
    xbm = (dna)? _dbval : _pbval;

    endgaps = 0;                                /* 1 to penalize endgaps */
    ofile = "align.out";                        /* output file */

    nw( );                                /* fill in the matrix, get the possible jumps */
    readjumps( );                            /* get the actual jumps */
    print( );                                /* print stats, alignment */

    cleanup(0);                                /* unlink any tmp files */
}

```

**Table 1 (cont')**

```

/* do the alignment, return best score: main( )
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */
nw( )
{
    char      *px, *py;           /* seqs and ptrs */
    int       *ndely, *dely;      /* keep track of dely */
    int       ndelx, delx;        /* keep track of delx */
    int       *tmp;              /* for swapping row0, row1 */
    int       mis;               /* score for each type */
    int       ins0, ins1;         /* insertion penalties */
    register  id;                /* diagonal index */
    register  ij;               /* jmp index */
    register  *col0, *col1;       /* score for curr, last row */
    register  xx, yy;            /* index into seqs */

    dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

    ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
    dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
    col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
    col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
    ins0 = (dna)? DINS0 : PINS0;
    ins1 = (dna)? DINS1 : PINS1;

    smax = -10000;
    if (endgaps) {
        for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
            col0[yy] = dely[yy] = col0[yy-1] - ins1;
            ndely[yy] = yy;
        }
        col0[0] = 0;          /* Waterman Bull Math Biol 84 */
    }
    else
        for (yy = 1; yy <= len1; yy++)
            dely[yy] = -ins0;

    /* fill in match matrix
     */
    for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
         */
        if (endgaps) {
            if (xx == 1)
                col1[0] = delx = -(ins0+ins1);
            else
                col1[0] = delx = col0[0] - ins1;
            ndelx = xx;
        }
        else {
            col1[0] = 0;
            delx = -ins0;
            ndelx = 0;
        }
    }
}

```

nw



Table 1 (cont')

...nw

```

for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
    mis = col0[yy-1];
    if (dna)
        mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
    else
        mis += _day[*px-'A'][*py-'A'];

    /* update penalty for del in x seq;
     * favor new del over ongong del
     * ignore MAXGAP if weighting endgaps
     */
    if (endgaps || ndely[yy] < MAXGAP) {
        if (col0[yy] - ins0 >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0 + ins1);
            ndely[yy] = 1;
        } else {
            dely[yy] -= ins1;
            ndely[yy]++;
        }
    } else {
        if (col0[yy] - (ins0 + ins1) >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0 + ins1);
            ndely[yy] = 1;
        } else
            ndely[yy]++;
    }

    /* update penalty for del in y seq;
     * favor new del over ongong del
     */
    if (endgaps || ndelx < MAXGAP) {
        if (col1[yy-1] - ins0 >= delx) {
            delx = col1[yy-1] - (ins0 + ins1);
            ndelx = 1;
        } else {
            delx -= ins1;
            ndelx++;
        }
    } else {
        if (col1[yy-1] - (ins0 + ins1) >= delx) {
            delx = col1[yy-1] - (ins0 + ins1);
            ndelx = 1;
        } else
            ndelx++;
    }

    /* pick the maximum score; we're favoring
     * mis over any del and delx over dely
     */

```

Table 1 (cont')

...nw

```

id = xx - yy + len1 - 1;
if (mis >= delx && mis >= dely[yy])
    coll[yy] = mis;
else if (delx >= dely[yy]) {
    coll[yy] = delx;
    ij = dx[id].ijmp;
    if (dx[id].jp.n[0] && (!dna || (ndelx >= MAXJMP
    && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
        dx[id].ijmp++;
        if (+ij >= MAXJMP) {
            writejumps(id);
            ij = dx[id].ijmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
        }
        dx[id].jp.n[ij] = ndelx;
        dx[id].jp.x[ij] = xx;
        dx[id].score = delx;
    }
} else {
    coll[yy] = dely[yy];
    ij = dx[id].ijmp;
    if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
    && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
        dx[id].ijmp++;
        if (+ij >= MAXJMP) {
            writejumps(id);
            ij = dx[id].ijmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
        }
        dx[id].jp.n[ij] = -ndely[yy];
        dx[id].jp.x[ij] = xx;
        dx[id].score = dely[yy];
    }
}
if (xx == len0 && yy < len1) {
    /* last col
    */
    if (endgaps)
        coll[yy] -= ins0+ins1*(len1-yy);
    if (coll[yy] > smax) {
        smax = coll[yy];
        dmax = id;
    }
}
}
if (endgaps && xx < len0)
    coll[yy-1] -= ins0+ins1*(len0-xx);
if (coll[yy-1] > smax) {
    smax = coll[yy-1];
    dmax = id;
}
tmp = col0; col0 = coll; coll = tmp;
}
(void) free((char *)ndely);
(void) free((char *)dely);
(void) free((char *)col0);
(void) free((char *)coll);
}

```

Table 1 (cont')

```

/*
 *
 * print( ) -- only routine visible outside this module
 *
 * static:
 * getmat( ) -- trace back best path, count matches: print( )
 * pr_align( ) -- print alignment of described in array p[]: print( )
 * dumpblock( ) -- dump a block of lines with numbers, stars: pr_align( )
 * nums( ) -- put out a number line: dumpblock( )
 * putline( ) -- put out a line (name, [num], seq, [num]): dumpblock( )
 * stars( ) -- put a line of stars: dumpblock( )
 * stripname( ) -- strip any path and prefix from a seqname
 */

#include "nw.h"

#define SPC      3
#define P_LINE  256    /* maximum output line */
#define P_SPC    3      /* space between name or num and seq */

extern _day[26][26];
int      olen;          /* set output line length */
FILE     *fx;           /* output file */

print( )
{
    int      lx, ly, firstgap, lastgap;    /* overlap */

    if ((fx = fopen(ofile, "w")) == 0) {
        fprintf(stderr, "%s: can't write %s\n", prog, ofile);
        cleanup(1);
    }
    fprintf(fx, "<first sequence: %s (length = %d)\n", namex[0], len0);
    fprintf(fx, "<second sequence: %s (length = %d)\n", namex[1], len1);
    olen = 60;
    lx = len0;
    ly = len1;
    firstgap = lastgap = 0;
    if (dmax < len1 - 1) { /* leading gap in x */
        pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
    else if (dmax > len1 - 1) { /* leading gap in y */
        pp[1].spc = firstgap = dmax - (len1 - 1);
        lx -= pp[1].spc;
    }
    if (dmax0 < len0 - 1) { /* trailing gap in x */
        lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
    }
    else if (dmax0 > len0 - 1) { /* trailing gap in y */
        lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
    getmat(lx, ly, firstgap, lastgap);
    pr_align();
}

```

print

Table 1 (cont')

```

/*
 * trace back the best path, count matches
 */
static
getmat(lx, ly, firstgap, lastgap)
    int      lx, ly;
    int      firstgap, lastgap;
{
    int      nm, i0, i1, siz0, siz1;
    char     outx[32];
    double   pct;
    register n0, n1;
    register char *p0, *p1;

    /* get total matches, score
     */
    i0 = i1 = siz0 = siz1 = 0;
    p0 = seqx[0] + pp[1].spc;
    p1 = seqx[1] + pp[0].spc;
    n0 = pp[1].spc + 1;
    n1 = pp[0].spc + 1;

    nm = 0;
    while ( *p0 && *p1 ) {
        if (siz0) {
            p1++;
            n1++;
            siz0--;
        }
        else if (siz1) {
            p0++;
            n0++;
            siz1--;
        }
        else {
            if (xbm[*p0-'A']&xbm[*p1-'A'])
                nm++;
            if (n0++ == pp[0].x[i0])
                siz0 = pp[0].n[i0++];
            if (n1++ == pp[1].x[i1])
                siz1 = pp[1].n[i1++];
            p0++;
            p1++;
        }
    }

    /* pct homology:
     * if penalizing endgaps, base is the shorter seq
     * else, knock off overhangs and take shorter core
     */
    if (endgaps)
        lx = (len0 < len1)? len0 : len1;
    else
        lx = (lx < ly)? lx : ly;
    pct = 100.*(double)nm/(double)lx;
    fprintf(fx, "\n");
    fprintf(fx, "< %d match%s in an overlap of %d: %.2f percent similarity\n",
        nm, (nm == 1)? "" : "es", lx, pct);
}

```

getmat

**Table 1 (cont')**

```

fprintf(fx, "< gaps in first sequence: %d", gapx);
if (gapx) {
    (void) sprintf(outx, " (%d %s%s)",
        ngapx, (dna)? "base": "residue", (ngapx == 1)? "": "s");
    fprintf(fx, "%s", outx);

    fprintf(fx, ", gaps in second sequence: %d", gapy);
    if (gapy) {
        (void) sprintf(outx, " (%d %s%s)",
            ngapy, (dna)? "base": "residue", (ngapy == 1)? "": "s");
        fprintf(fx, "%s", outx);
    }
    if (dna)
        fprintf(fx,
            "\n< score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
            smax, DMAT, DMIS, DINSO, DINSI);
    else
        fprintf(fx,
            "\n< score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
            smax, PINSO, PINSI);
    if (endgaps)
        fprintf(fx,
            "< endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
            firstgap, (dna)? "base": "residue", (firstgap == 1)? "": "s",
            lastgap, (dna)? "base": "residue", (lastgap == 1)? "": "s");
    else
        fprintf(fx, "< endgaps not penalized\n");
}

static      nm;          /* matches in core -- for checking */
static      lmax;        /* lengths of stripped file names */
static      ij[2];       /* jmp index for a path */
static      nc[2];       /* number at start of current line */
static      ni[2];       /* current elem number -- for gapping */
static      siz[2];
static char *ps[2];      /* ptr to current element */
static char *po[2];      /* ptr to next output char slot */
static char out[2][P_LINE]; /* output line */
static char star[P_LINE]; /* set by stars() */

/*
 * print alignment of described in struct path pp[]
 */
static
pr_align()
{
    int      nn;          /* char count */
    int      more;
    register i;

    for (i = 0, lmax = 0; i < 2; i++) {
        nn = stripname(namex[i]);
        if (nn > lmax)
            lmax = nn;

        nc[i] = 1;
        ni[i] = 1;
        siz[i] = ij[i] = 0;
        ps[i] = seqx[i];
        po[i] = out[i];
    }
}

```

...getmat

pr\_align

Table 1 (cont')

```

for (nn = nm = 0, more = 1; more; ) {
    for (i = more = 0; i < 2; i++) {
        /*
         * do we have more of this sequence?
         */
        if (!*ps[i])
            continue;

        more++;

        if (pp[i].spc) { /* leading space */
            *po[i]++ = ' ';
            pp[i].spc--;
        }
        else if (siz[i]) { /* in a gap */
            *po[i]++ = '-';
            siz[i]--;
        }
        else { /* we're putting a seq element
            */
            *po[i] = *ps[i];
            if (islower(*ps[i]))
                *ps[i] = toupper(*ps[i]);
            po[i]++;
            ps[i]++;

            /*
             * are we at next gap for this seq?
             */
            if (ni[i] == pp[i].x[ij[i]]) {
                /*
                 * we need to merge all gaps
                 * at this location
                 */
                siz[i] = pp[i].n[ij[i]] + +;
                while (ni[i] == pp[i].x[ij[i]])
                    siz[i] += pp[i].n[ij[i]] + +;
            }
            ni[i]++;
        }
    }
    if (++nn == olen || !more && nn) {
        dumpblock();
        for (i = 0; i < 2; i++)
            po[i] = out[i];
        nn = 0;
    }
}

/*
 * dump a block of lines, including numbers, stars: pr_align( )
 */
static
dumpblock( )
{
    register i;

    for (i = 0; i < 2; i++)
        *po[i] = '\0';
}

```

...pr\_align

dumpblock

Table 1 (cont')

...dumpblock

```

(void) putc('\n', fx);
for (i = 0; i < 2; i++) {
    if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
        if (i == 0)
            nums(i);
        if (i == 0 && *out[1])
            stars( );
        putline(i);
        if (i == 0 && *out[1])
            fprintf(fx, star);
        if (i == 1)
            nums(i);
    }
}

/*
 * put out a number line: dumpblock( )
 */
static
nums(ix)
int    ix;    /* index in out[] holding seq line */
{
    char    nline[P_LINE];
    register i, j;
    register char    *pn, *px, *py;

    for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
        *pn = ' ';
    for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
        if (*py == ' ' || *py == '-')
            *pn = ' ';
        else {
            if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                j = (i < 0)? -i : i;
                for (px = pn; j /= 10, px--)
                    *px = j%10 + '0';
                if (i < 0)
                    *px = '-';
            }
            else
                *pn = ' ';
            i++;
        }
    }
    *pn = '\0';
    nc[ix] = i;
    for (pn = nline; *pn; pn++)
        (void) putc(*pn, fx);
    (void) putc('\n', fx);
}

/*
 * put out a line (name, [num], seq, [num]): dumpblock( )
 */
static
putline(ix)
int    ix;
{
    putline

```

Table 1 (cont')

```

int          i;
register char *px;

for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
    (void) putc(*px, fx);
for (; i < lmax+P_SPC; i++)
    (void) putc(' ', fx);

/* these count from 1:
 * ni[] is current element (from 1)
 * nc[] is number at start of current line
 */
for (px = out[ix]; *px; px++)
    (void) putc(*px&0x7F, fx);
(void) putc('\n', fx);
}

/*
 * put a line of stars (seqs always in out[0], out[1]): dumpblock( )
 */
static
stars( )
{
    int          i;
    register char *p0, *p1, cx, *px;

    if (!(*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
        !(*out[1] || (*out[1] == ' ' && *(po[1]) == ' ')))
        return;
    px = star;
    for (i = lmax+P_SPC; i; i--)
        *px++ = ' ';

    for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
        if (isalpha(*p0) && isalpha(*p1)) {
            if (xbm[*p0-'A']&xbm[*p1-'A']) {
                cx = '*';
                nm++;
            }
            else if (!dna && _day[*p0-'A'][*p1-'A'] > 0)
                cx = '.';
            else
                cx = ' ';
        }
        else
            cx = ' ';
        *px++ = cx;
    }
    *px++ = '\n';
    *px = '\0';
}

```

...putline

stars



**Table 1 (cont')**

```

/*
 * strip path or prefix from pn, return len: pr_align( )
 */
static
stripname(pn)
    char    *pn;    /* file name (may be path) */
{
    register char    *px, *py;

    py = 0;
    for (px = pn; *px; px++)
        if (*px == '/')
            py = px + 1;
    if (py)
        (void) strcpy(pn, py);
    return(strlen(pn));
}

```

**stripname**

**Table 1 (cont')**

```

/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * g_calloc() -- calloc() with error checkin
 * readjumps() -- get the good jumps, from tmp file if necessary
 * writejumps() -- write a filled array of jumps to a tmp file: nw()
 */
#include "nw.h"
#include <sys/file.h>

char    *jname = "/tmp/homgXXXXXX";          /* tmp file for jumps */
FILE    *fj;

int      cleanup();                          /* cleanup tmp file */
long     lseek();

/*
 * remove any tmp file if we blow
 */
cleanup(i)
    int    i;
{
    if (fj)
        (void) unlink(jname);
    exit(i);
}

/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
char    *
getseq(file, len)
    char    *file;    /* file name */
    int     *len;     /* seq len */
{
    char    line[1024], *pseq;
    register char    *px, *py;
    int     natgc, tlen;
    FILE    *fp;

    if ((fp = fopen(file, "r")) == 0) {
        fprintf(stderr, "%s: can't read %s\n", prog, file);
        exit(1);
    }
    tlen = natgc = 0;
    while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++)
            if (isupper(*px) || islower(*px))
                tlen++;
    }
    if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr, "%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
    }
    pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
}

```

cleanup

getseq

Table 1 (cont')

...getseq

```

py = pseq + 4;
*len = tlen;
rewind(fp);

while (fgets(line, 1024, fp)) {
    if (*line == ';' || *line == '<' || *line == '>')
        continue;
    for (px = line; *px != '\n'; px++) {
        if (isupper(*px))
            *py++ = *px;
        else if (islower(*px))
            *py++ = toupper(*px);
        if (index("ATGCU", *(py-1)))
            natgc++;
    }
    *py++ = '\0';
    *py = '\0';
    (void) fclose(fp);
    dna = natgc > (tlen/3);
    return(pseq+4);
}

```

```

char *
g_calloc(msg, nx, sz)
char *msg;          /* program, calling routine */
int nx, sz;         /* number and size of elements */
{
    char *px, *calloc();

    if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
        if (*msg) {
            fprintf(stderr, "%s: g_calloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
            exit(1);
        }
    }
    return(px);
}

```

g\_calloc

```

/*
 * get final jmps from dx[] or tmp file, set pp[], reset dmax: main( )
 */

```

```

readjmps()
{
    int fd = -1;
    int siz, i0, i1;
    register i, j, xx;

    if (fj) {
        (void) fclose(fj);
        if ((fd = open(jname, O_RDONLY, 0)) < 0) {
            fprintf(stderr, "%s: can't open( ) %s\n", prog, jname);
            cleanup(1);
        }
    }
    for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; i++) {
        while (1) {
            for (j = dx[dmax].ijmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                ;
        }
    }
}

```

readjmps

**Table 1 (cont')**

...readjumps

```

    if (j < 0 && dx[dmax].offset && fj) {
        (void) lseek(fd, dx[dmax].offset, 0);
        (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
        (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
        dx[dmax].ijmp = MAXJMP-1;
    }
    else
        break;
}
if (i >= JMPS) {
    fprintf(stderr, "%s: too many gaps in alignment\n", prog);
    cleanup(1);
}
if (j >= 0) {
    siz = dx[dmax].jp.n[j];
    xx = dx[dmax].jp.x[j];
    dmax += siz;
    if (siz < 0) { /* gap in second seq */
        pp[1].n[i1] = -siz;
        xx += siz;
        /* id = xx - yy + len1 - 1 */
        /*
        pp[1].x[i1] = xx - dmax + len1 - 1;
        gapy++;
        ngapy -= siz;
        */
        /* ignore MAXGAP when doing endgaps */
        siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
        i1++;
    }
    else if (siz > 0) { /* gap in first seq */
        pp[0].n[i0] = siz;
        pp[0].x[i0] = xx;
        gapx++;
        ngapx += siz;
        /* ignore MAXGAP when doing endgaps */
        siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
        i0++;
    }
}
else
    break;
}

/* reverse the order of jumps */
*/
for (j = 0, i0--; j < i0; j++, i0--) {
    i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
    i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
}
for (j = 0, i1--; j < i1; j++, i1--) {
    i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
    i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
}
if (fd >= 0)
    (void) close(fd);
if (fj) {
    (void) unlink(jname);
    fj = 0;
    offset = 0;
}
}

```

**Table 1 (cont')**

```

/*
 * write a filled jmp struct offset of the prev one (if any): nw( )
 */
writejumps(ix)                                     writejumps
{
    int      ix;
    char      *mktemp( );

    if (!fj) {
        if (mktemp(jname) < 0) {
            fprintf(stderr, "%s: can't mktemp( ) %s\n", prog, jname);
            cleanup(1);
        }
        if ((fj = fopen(jname, "w")) == 0) {
            fprintf(stderr, "%s: can't write %s\n", prog, jname);
            exit(1);
        }
    }
    (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
    (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
}

```

Table 2

PRO	XXXXXXXXXXXXXXXXXX	(Length = 15 amino acids)
Comparison Protein	XXXXXXYYYYYYY	(Length = 12 amino acids)

5      % amino acid sequence identity =  
 (the number of identically matching amino acid residues between the two polypeptide sequences as determined by  
 ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =  
 5 divided by 15 = 33.3%

10      Table 3

PRO	XXXXXXXXXXXX	(Length = 10 amino acids)
Comparison Protein	XXXXXXYYYYYYZZYZ	(Length = 15 amino acids)

15      % amino acid sequence identity =  
 (the number of identically matching amino acid residues between the two polypeptide sequences as determined by  
 ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =  
 5 divided by 10 = 50%

20      Table 4

PRO-DNA	NNNNNNNNNNNNNN	(Length = 14 nucleotides)
Comparison DNA	NNNNNNLLLLLLLLLL	(Length = 16 nucleotides)

25      % nucleic acid sequence identity =  
 (the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-  
 2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =  
 6 divided by 14 = 42.9%

30      Table 5

PRO-DNA	NNNNNNNNNNNN	(Length = 12 nucleotides)
Comparison DNA	NNNNLLL VV	(Length = 9 nucleotides)

35      % nucleic acid sequence identity =  
 (the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-  
 2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =  
 4 divided by 12 = 33.3%

## 5.2. Compositions and Methods of the Invention

### 5.2.1. PRO Variants

5 In addition to the full-length native sequence PRO polypeptides described herein, it is contemplated that PRO variants can be prepared. PRO variants can be prepared by introducing appropriate nucleotide changes into the PRO DNA, and/or by synthesis of the desired PRO polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the PRO polypeptide such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

10 Variations in the native full-length sequence PRO polypeptide or in various domains of the PRO polypeptide described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the PRO polypeptide that results in a change in the amino acid sequence of the PRO polypeptide as compared with the native sequence PRO polypeptide. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the PRO polypeptide. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the PRO polypeptide with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, *i.e.*, conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

20 In particular embodiments, conservative substitutions of interest are shown in Table 6 under the heading of preferred substitutions. If such substitutions result in a change in biological activity, then more substantial changes, denominated exemplary substitutions in Table 6, or as further described below in reference to amino acid classes, are introduced and the products screened.

Table 6

	Original <u>Residue</u>	Exemplary <u>Substitutions</u>	Preferred <u>Substitutions</u>
	Ala (A)	val; leu; ile	val
5	Arg (R)	lys; gln; asn	lys
	Asn (N)	gln; his; lys; arg	gln
	Asp (D)	glu	glu
	Cys (C)	ser	ser
	Gln (Q)	asn	asn
10	Glu (E)	asp	asp
	Gly (G)	pro; ala	ala
	His (H)	asn; gln; lys; arg	arg
	Ile (I)	leu; val; met; ala; phe; norleucine	leu
15	Leu (L)	norleucine; ile; val; met; ala; phe	ile
	Lys (K)	arg; gln; asn	arg
	Met (M)	leu; phe; ile	leu
	Phe (F)	leu; val; ile; ala; tyr	leu
20	Pro (P)	ala	ala
	Ser (S)	thr	thr
	Thr (T)	ser	ser
	Trp (W)	tyr; phe	tyr
	Tyr (Y)	trp; phe; thr; ser	phe
25	Val (V)	ile; leu; met; phe; ala; norleucine	leu

Substantial modifications in function or immunological identity of the PRO polypeptide are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

- (1) hydrophobic: norleucine, met, ala, val, leu, ile;
- (2) neutral hydrophilic: cys, ser, thr;
- (3) acidic: asp, glu;
- 35 (4) basic: asn, gln, his, lys, arg;
- (5) residues that influence chain orientation: gly, pro; and
- (6) aromatic: trp, tyr, phe.

Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter *et al.*, Nucl. Acids Res., 13:4331 (1986); Zoller *et al.*, Nucl. Acids Res., 10:6487 (1987)], cassette mutagenesis [Wells *et al.*, Gene, 34:315 (1985)], restriction selection mutagenesis [Wells *et al.*, Philos. Trans. R. Soc. London SerA, 317:415



(1986)] or other known techniques can be performed on the cloned DNA to produce the PRO variant DNA.

Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant [Cunningham and Wells, Science, 244: 1081-1085 (1989)]. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

#### 5.2.2. Modifications of PRO Polypeptides

Covalent modifications of PRO polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C- terminal residues of the PRO polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking the PRO polypeptide to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO antibodies, and vice-versa. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimide.

Other modifications include deamidation of glutamyl and asparaginy residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the  $\alpha$ -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the PRO polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in the native sequence PRO polypeptide (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the native sequence PRO polypeptide. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

Addition of glycosylation sites to the PRO polypeptide may be accomplished by altering the amino acid sequence. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO polypeptide (for O-linked glycosylation sites). The PRO amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO polypeptide at preselected bases such that codons are generated that will translate into the desired amino

acids.

Another means of increasing the number of carbohydrate moieties on the PRO polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, *e.g.*, in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

5 Removal of carbohydrate moieties present on the PRO polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, *et al.*, Arch. Biochem. Biophys., 259:52 (1987) and by Edge *et al.*, Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and  
10 exo-glycosidases as described by Thotakura *et al.*, Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of the PRO polypeptide comprises linking the PRO polypeptide to one of a variety of nonproteinaceous polymers, *e.g.*, polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

15 The PRO polypeptide of the present invention may also be modified in a way to form a chimeric molecule comprising the PRO polypeptide fused to another, heterologous polypeptide or amino acid sequence.

In one embodiment, such a chimeric molecule comprises a fusion of the PRO polypeptide with a protein transduction domain which targets the PRO polypeptide for delivery to various tissues and more particularly across the brain blood barrier, using, for example, the protein transduction domain of human immunodeficiency virus TAT  
20 protein (Schwarze *et al.*, 1999, *Science* 285: 1569-72).

In another embodiment, such a chimeric molecule comprises a fusion of the PRO polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the PRO polypeptide. The presence of such epitope-tagged forms of the PRO polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of  
25 the epitope tag enables the PRO polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-His) or poly-histidine-glycine (poly-His-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field *et al.*, Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan *et al.*, Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky *et al.*, Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp  
30 *et al.*, BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin *et al.*, Science, 255:192-194 (1992)]; an  $\alpha$ -tubulin epitope peptide [Skinner *et al.*, J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth *et al.*, Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

35 In an alternative embodiment, the chimeric molecule may comprise a fusion of the PRO polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions

preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of a PRO polypeptide in place of at least one variable region within an Ig molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH2 and CH3, or the hinge, CH1, CH2 and CH3 regions of an IgG1 molecule. For the production of immunoglobulin fusions *see also*, U.S. Patent No. 5,428,130 issued June 27, 1995.

5                   5.2.3. Preparation of the PRO Polypeptide

10           The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO polypeptides. In particular, cDNAs encoding PRO polypeptides have been identified and isolated, as disclosed in further detail in the Examples below. It is noted that proteins produced in separate expression rounds may be given different PRO numbers but the UNQ number is unique for any given DNA and the encoded protein, and will not be changed. However, for sake of simplicity, in the present specification the protein encoded by the PRO DNA as well as all further native homologues and variants included in the foregoing definition of PRO polypeptides, will be referred to as "PRO" regardless of their origin or mode of preparation.

15           The description below relates primarily to production of PRO polypeptides by culturing cells transformed or transfected with a vector containing nucleic acid encoding PRO polypeptides. It is, of course, contemplated that alternative methods that are well known in the art may be employed to prepare the PRO polypeptide. For instance, the PRO polypeptide sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques. *See, e.g., Stewart et al., Solid-Phase Peptide Synthesis* (W.H. Freeman Co.: San Francisco, CA, 1969); Merrifield, *J. Am. Chem. Soc.*, **85**: 2149-2154 (1963). *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, with an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of the PRO polypeptide may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the full-length PRO polypeptide.

20                               5.2.3.1. Isolation of DNA Encoding PRO Polypeptides

25           DNA encoding the PRO polypeptide may be obtained from a cDNA library prepared from tissue believed to possess the mRNA encoding the PRO polypeptide and to express it at a detectable level. Accordingly, DNAs encoding the human PRO polypeptide can be conveniently obtained from cDNA libraries prepared from human tissues, such as described in the Examples. The gene encoding the PRO polypeptide may also be obtained from a genomic library or by oligonucleotide synthesis.

30           Libraries can be screened with probes (such as antibodies to the PRO polypeptide or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook *et al., supra*. An alternative means to isolate the gene encoding the PRO polypeptide is to use PCR methodology. Sambrook *et al., supra*; Dieffenbach *et al., PCR Primer: A Laboratory Manual* (New York: Cold

Spring Harbor Laboratory Press, 1995).

The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being  
5 screened. Methods of labeling are well known in the art, and include the use of radiolabels like <sup>32</sup>P-labeled ATP, biotinylation, or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook *et al.*, *supra*.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases.  
10 Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined through sequence alignment using computer software programs such as ALIGN, DNASTar, and INHERIT, which employ various algorithms to measure homology.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using  
15 conventional primer extension procedures as described in Sambrook *et al.*, *supra*, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

#### 5.2.3.2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for PRO polypeptide production and cultured in conventional nutrient media modified as appropriate for inducing promoters,  
20 selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH, and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: A Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook *et al.*, *supra*.

25 Methods of transfection are known to the ordinarily skilled artisan, for example, CaPO<sub>4</sub> treatment and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook *et al.*, *supra*, or electroporation is generally used for prokaryotes or other cells that contain substantial cell-wall barriers. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw *et al.*, Gene,  
30 23: 315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transformations have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen *et al.*, J. Bact., 130:  
946 (1977) and Hsiao *et al.*, Proc. Natl. Acad. Sci. (USA), 76: 3829 (1979). However, other methods for  
35 introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene or polyornithine, may also be used. For various techniques for

transforming mammalian cells, *see*, Keown *et al.*, Methods in Enzymology, 185: 527-537 (1990) and Mansour *et al.*, Nature, 336: 348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include, but are not limited to, eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325); and K5 772 (ATCC 53,635). Other suitable prokaryotic host cells include Enterobacteriaceae such as *Escherichia*, *e.g.*, *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, *e.g.*, *Salmonella typhimurium*, *Serratia*, *e.g.*, *Serratia marcescans*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (*e.g.*, *B. licheniformis* 41P disclosed in DD 266,710 published 12 April 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are illustrative rather than limiting. Strain W3110 is one particularly preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For example, strain W3110 may be modified to effect a genetic mutation in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype *tonA*; *E. coli* W3110 strain 9E4, which has the complete genotype *tonA ptr3*; *E. coli* W3110 strain 27C7 (ATCC 55,244), which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT kan'*; *E. coli* W3110 strain 37D6, which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7 ilvG kan'*; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vitro* methods of cloning, *e.g.*, PCR or other nucleic acid polymerase reactions, are suitable.

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for vectors encoding the PRO polypeptide. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, Nature, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Patent No. 4,943,529; Fleer *et al.*, Bio/Technology, 9: 968-975 (1991)) such as, *e.g.*, *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt *et al.*, J. Bacteriol., 737 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickerhamii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilae* (ATCC 36,906; Van den Berg *et al.*, Bio/Technology, 8: 135 (1990)), *K. thermotolerans*, and *K. marxianus*; *yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna *et al.*, J. Basic Microbiol., 28: 265-278 [1988]); *Candida*; *Trichoderma reesia* (EP 244,234); *Neurospora crassa* (Case *et al.*, Proc. Natl. Acad. Sci. USA, 76: 5259-5263 [1979]); *Schwanniomyces* such as *Schwanniomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, *e.g.*, *Neurospora*, *Penicillium*, *Tolypocladium* (WO 91/00357 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* (Ballance *et al.*, Biochem. Biophys. Res. Commun., 112: 284-289 [1983]; Tilburn *et al.*, Gene, 26: 205-221 [1983]; Yelton *et al.*, Proc. Natl. Acad. Sci. USA, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, EMBO J., 4: 475-479 [1985]). Methylophilic yeasts are suitable herein and include, but are not limited to, yeast capable of growth on methanol selected from the genera consisting of *Hansenula*, *Candida*, *Kloeckera*, *Pichia*, *Saccharomyces*,

*Torulopsis*, and *Rhodotorula*. A list of specific species that are exemplary of this class of yeasts may be found in C. Anthony, The Biochemistry of Methylootrophs, 269 (1982).

Suitable host cells for the expression of nucleic acid encoding glycosylated PRO polypeptides are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham *et al.*, J. Gen. Virol., 36: 59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

#### 5.2.3.3. Selection and Use of a Replicable Vector

The nucleic acid (*e.g.*, cDNA or genomic DNA) encoding the PRO polypeptide may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence if the sequence is to be secreted, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques that are known to the skilled artisan.

The PRO polypeptide may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the DNA encoding the PRO polypeptide that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, *e.g.*, the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces*  $\alpha$ -factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2 $\mu$  plasmid origin

is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV, or BPV) are useful for cloning vectors in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, *e.g.*, ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, *e.g.*, the gene encoding D-alanine racemase for *Bacilli*.

An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the nucleic acid encoding the PRO polypeptide such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub *et al.*, Proc. Natl. Acad. Sci. USA, 77: 4216 (1980). A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7. Stinchcomb *et al.*, Nature, 282: 39 (1979); Kingsman *et al.*, Gene, 7: 141 (1979); Tschemper *et al.*, Gene, 10: 157 (1980). The *trp1* gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1. Jones, Genetics, 85: 12 (1977).

Expression and cloning vectors usually contain a promoter operably linked to the nucleic acid sequence encoding the PRO polypeptide to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the  $\beta$ -lactamase and lactose promoter systems (Chang *et al.*, Nature, 275: 615 (1978); Goeddel *et al.*, Nature, 281: 544 (1979)), alkaline phosphatase, a tryptophan (*trp*) promoter system (Goeddel, Nucleic Acids Res., 8: 4057 (1980); EP 36,776), and hybrid promoters such as the *tac* promoter (deBoer *et al.*, Proc. Natl. Acad. Sci. USA, 80: 21-25 (1983)). Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding the PRO polypeptide.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase (Hitzeman *et al.*, J. Biol. Chem., 255: 2073 (1980)) or other glycolytic enzymes (Hess *et al.*, J. Adv. Enzyme Reg., 7: 149 (1968); Holland, Biochemistry, 17: 4900 (1978)), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters that are inducible promoters having the additional advantage of transcription controlled by growth conditions are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

PRO nucleic acid transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus, and Simian Virus 40 (SV40); by heterologous mammalian promoters, *e.g.*, the actin

promoter or an immunoglobulin promoter; and by heat-shock promoters, provided such promoters are compatible with the host cell systems.

Transcription of a DNA encoding the PRO polypeptide by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin,  $\alpha$ -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the sequence coding for PRO polypeptides, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding the PRO polypeptide.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of the PRO polypeptide in recombinant vertebrate cell culture are described in Gething *et al.*, Nature, 293: 620-625 (1981); Mantei *et al.*, Nature, 281: 40-46 (1979); EP 117,060; and EP 117,058.

#### 5.2.3.4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA (Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)), dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native-sequence PRO polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to DNA encoding the PRO polypeptide and encoding a specific antibody epitope.

#### 5.2.3.5. Purification of PRO Polypeptides

Forms of PRO polypeptides may be recovered from culture medium or from host cell lysates. If



membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g., TRITON-X™ 100) or by enzymatic cleavage. Cells employed in expression of nucleic acid encoding the PRO polypeptide can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell-lysing agents. It may be desired to purify the PRO polypeptide from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO polypeptide. Various methods of protein purification may be employed and such methods are known in the art and described, for example, in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice (Springer-Verlag: New York, 1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular PRO polypeptide produced.

#### 5.2.4. Uses of PRO Polypeptides

##### 5.2.4.1. Assays for Cardiovascular, Endothelial, and Angiogenic Activity

Various assays can be used to test the polypeptide herein for cardiovascular, endothelial, and angiogenic activity. Such assays include those provided in the Examples below.

Assays for testing for endothelin antagonist activity, as disclosed in U.S. Pat. No. 5,773,414, include a rat heart ventricle binding assay where the polypeptide is tested for its ability to inhibit iodinated endothelin-1 binding in a receptor assay, an endothelin receptor binding assay testing for intact cell binding of radiolabeled endothelin-1 using rabbit renal artery vascular smooth muscle cells, an inositol phosphate accumulation assay where functional activity is determined in Rat-1 cells by measuring intra-cellular levels of second messengers, an arachidonic acid release assay that measures the ability of added compounds to reduce endothelin-stimulated arachidonic acid release in cultured vascular smooth muscles, *in vitro* (isolated vessel) studies using endothelium from male New Zealand rabbits, and *in vivo* studies using male Sprague-Dawley rats.

Assays for tissue generation activity include, without limitation, those described in WO 95/16035 (bone, cartilage, tendon); WO 95/05846 (nerve, neuronal), and WO 91/07491 (skin, endothelium).

Assays for wound-healing activity include, for example, those described in Winter, Epidermal Wound Healing, Maibach, HI and Rovee, DT, eds. (Year Book Medical Publishers, Inc., Chicago), pp. 71-112, as modified by the article of Eaglstein and Mertz, J. Invest. Dermatol., 71: 382-384 (1978).

An assay to screen for a test molecule relating to a PRO polypeptide that binds an endothelin B<sub>1</sub> (ETB<sub>1</sub>) receptor polypeptide and modulates signal transduction activity involves providing a host cell transformed with a DNA encoding endothelin B<sub>1</sub> receptor polypeptide, exposing the cells to the test candidate, and measuring endothelin B<sub>1</sub> receptor signal transduction activity, as described, e.g., in U.S. Pat. No. 5,773,223.

There are several cardiac hypertrophy assays. *In vitro* assays include induction of spreading of adult rat cardiac myocytes. In this assay, ventricular myocytes are isolated from a single (male Sprague-Dawley) rat,

essentially following a modification of the procedure described in detail by Piper *et al.*, "Adult ventricular rat heart muscle cells" in Cell Culture Techniques in Heart and Vessel Research, H.M. Piper, ed. (Berlin: Springer-Verlag, 1990), pp. 36-60. This procedure permits the isolation of adult ventricular myocytes and the long-term culture of these cells in the rod-shaped phenotype. Phenylephrine and Prostaglandin  $F_{2\alpha}$  ( $PGF_{2\alpha}$ ) have been shown to induce a spreading response in these adult cells. The inhibition of myocyte spreading induced by  $PGF_{2\alpha}$  or  $PGF_{2\alpha}$  analogs (e.g., fluprostenol) and phenylephrine by various potential inhibitors of cardiac hypertrophy is then tested.

One example of an *in vivo* assay is a test for inhibiting cardiac hypertrophy induced by fluprostenol *in vivo*. This pharmacological model tests the ability of the PRO polypeptide to inhibit cardiac hypertrophy induced in rats (e.g., male Wistar or Sprague-Dawley) by subcutaneous injection of fluprostenol (an agonist analog of  $PGF_{2\alpha}$ ). It is known that rats with pathologic cardiac hypertrophy induced by myocardial infarction have chronically elevated levels of extractable  $PGF_{2\alpha}$  in their myocardium. Lai *et al.*, Am. J. Physiol. (Heart Circ. Physiol.), 271: H2197-H2208 (1996). Accordingly, factors that can inhibit the effects of fluprostenol on myocardial growth *in vivo* are potentially useful for treating cardiac hypertrophy. The effects of the PRO polypeptide on cardiac hypertrophy are determined by measuring the weight of heart, ventricles, and left ventricle (normalized by body weight) relative to fluprostenol-treated rats not receiving the PRO polypeptide.

Another example of an *in vivo* assay is the pressure-overload cardiac hypertrophy assay. For *in vivo* testing it is common to induce pressure-overload cardiac hypertrophy by constriction of the abdominal aorta of test animals. In a typical protocol, rats (e.g., male Wistar or Sprague-Dawley) are treated under anesthesia, and the abdominal aorta of each rat is narrowed down just below the diaphragm. Beznak M., Can. J. Biochem. Physiol., 33: 985-94 (1955). The aorta is exposed through a surgical incision, and a blunted needle is placed next to the vessel. The aorta is constricted with a ligature of silk thread around the needle, which is immediately removed and which reduces the lumen of the aorta to the diameter of the needle. This approach is described, for example, in Rossi *et al.*, Am. Heart J., 124: 700-709 (1992) and O'Rourke and Reibel, P.S.E.M.B., 200: 95-100 (1992).

In yet another *in vivo* assay, the effect on cardiac hypertrophy following experimentally induced myocardial infarction (MI) is measured. Acute MI is induced in rats by left coronary artery ligation and confirmed by electrocardiographic examination. A sham-operated group of animals is also prepared as control animals. Earlier data have shown that cardiac hypertrophy is present in the group of animals with MI, as evidenced by an 18% increase in heart weight-to-body weight ratio. Lai *et al.*, *supra*. Treatment of these animals with candidate blockers of cardiac hypertrophy, e.g., the PRO polypeptide, provides valuable information about the therapeutic potential of the candidates tested. One further such assay test for induction of cardiac hypertrophy is disclosed in U.S. Pat. No. 5,773,415, using Sprague-Dawley rats.

For cancer, a variety of well-known animal models can be used to further understand the role of the genes identified herein in the development and pathogenesis of tumors, and to test the efficacy of candidate therapeutic agents, including antibodies and other antagonists of native PRO polypeptides, such as small-molecule antagonists. The *in vivo* nature of such models makes them particularly predictive of responses in human patients. Animal models of tumors and cancers (e.g., breast cancer, colon cancer, prostate cancer, lung cancer, etc.) include both non-recombinant and recombinant (transgenic) animals. Non-recombinant animal models include, for example,

rodent, *e.g.*, murine models. Such models can be generated by introducing tumor cells into syngeneic mice using standard techniques, *e.g.*, subcutaneous injection, tail vein injection, spleen implantation, intraperitoneal implantation, implantation under the renal capsule, or orthopin implantation, *e.g.*, colon cancer cells implanted in colonic tissue. *See, e.g.*, PCT publication No. WO 97/33551, published September 18, 1997. Probably the most often used animal species in oncological studies are immunodeficient mice and, in particular, nude mice. The observation that the nude mouse with thymic hypo/aplasia could successfully act as a host for human tumor xenografts has lead to its widespread use for this purpose. The autosomal recessive *nu* gene has been introduced into a very large number of distinct congenic strains of nude mouse, including, for example, ASW, A/He, AKR, BALB/c, B10.LP, C17, C3H, C57BL, C57, CBA, DBA, DDD, I/st, NC, NFR, NFS, NFS/N, NZB, NZC, NZW, P, RIII, and SJL. In addition, a wide variety of other animals with inherited immunological defects other than the nude mouse have been bred and used as recipients of tumor xenografts. For further details *see, e.g.*, The Nude Mouse in Oncology Research, E. Boven and B. Winograd, eds. (CRC Press, Inc., 1991).

The cells introduced into such animals can be derived from known tumor/cancer cell lines, such as any of the above-listed tumor cell lines, and, for example, the B104-1-1 cell line (stable NIH-3T3 cell line transfected with the *neu* protooncogene); *ras*-transfected NIH-3T3 cells; Caco-2 (ATCC HTB-37); or a moderately well-differentiated grade II human colon adenocarcinoma cell line, HT-29 (ATCC HTB-38); or from tumors and cancers. Samples of tumor or cancer cells can be obtained from patients undergoing surgery, using standard conditions involving freezing and storing in liquid nitrogen. Karmali *et al.*, Br. J. Cancer, 48: 689-696 (1983).

Tumor cells can be introduced into animals such as nude mice by a variety of procedures. The subcutaneous (s.c.) space in mice is very suitable for tumor implantation. Tumors can be transplanted s.c. as solid blocks, as needle biopsies by use of a trochar, or as cell suspensions. For solid-block or trochar implantation, tumor tissue fragments of suitable size are introduced into the s.c. space. Cell suspensions are freshly prepared from primary tumors or stable tumor cell lines, and injected subcutaneously. Tumor cells can also be injected as subdermal implants. In this location, the inoculum is deposited between the lower part of the dermal connective tissue and the s.c. tissue.

Animal models of breast cancer can be generated, for example, by implanting rat neuroblastoma cells (from which the *neu* oncogene was initially isolated), or *neu*-transformed NIH-3T3 cells into nude mice, essentially as described by Drebin *et al.* Proc. Nat. Acad. Sci. USA, 83: 9129-9133 (1986).

Similarly, animal models of colon cancer can be generated by passaging colon cancer cells in animals, *e.g.*, nude mice, leading to the appearance of tumors in these animals. An orthotopic transplant model of human colon cancer in nude mice has been described, for example, by Wang *et al.*, Cancer Research, 54: 4726-4728 (1994) and Too *et al.*, Cancer Research, 55: 681-684 (1995). This model is based on the so-called "METAMOUSE<sup>TM</sup>" sold by AntiCancer, Inc., (San Diego, California).

Tumors that arise in animals can be removed and cultured *in vitro*. Cells from the *in vitro* cultures can then be passaged to animals. Such tumors can serve as targets for further testing or drug screening. Alternatively, the tumors resulting from the passage can be isolated and RNA from pre-passage cells and cells isolated after one or more rounds of passage analyzed for differential expression of genes of interest. Such passaging techniques can

be performed with any known tumor or cancer cell lines.

For example, Meth A, CMS4, CMS5, CMS21, and WEHI-164 are chemically induced fibrosarcomas of BALB/c female mice (DeLeo *et al.*, J. Exp. Med., **146**: 720 (1977)), which provide a highly controllable model system for studying the anti-tumor activities of various agents. Palladino *et al.*, J. Immunol., **138**: 4023-4032 (1987). Briefly, tumor cells are propagated *in vitro* in cell culture. Prior to injection into the animals, the cell lines are washed and suspended in buffer, at a cell density of about  $10 \times 10^6$  to  $10 \times 10^7$  cells/ml. The animals are then infected subcutaneously with 10 to 100  $\mu$ l of the cell suspension, allowing one to three weeks for a tumor to appear.

In addition, the Lewis lung (3LL) carcinoma of mice, which is one of the most thoroughly studied experimental tumors, can be used as an investigational tumor model. Efficacy in this tumor model has been correlated with beneficial effects in the treatment of human patients diagnosed with small-cell carcinoma of the lung (SCCL). This tumor can be introduced in normal mice upon injection of tumor fragments from an affected mouse or of cells maintained in culture. Zupi *et al.*, Br. J. Cancer, **41**: suppl. 4, 30 (1980). Evidence indicates that tumors can be started from injection of even a single cell and that a very high proportion of infected tumor cells survive. For further information about this tumor model see, Zacharski, Haemostasis, **16**: 300-320 (1986).

One way of evaluating the efficacy of a test compound in an animal model with an implanted tumor is to measure the size of the tumor before and after treatment. Traditionally, the size of implanted tumors has been measured with a slide caliper in two or three dimensions. The measure limited to two dimensions does not accurately reflect the size of the tumor; therefore, it is usually converted into the corresponding volume by using a mathematical formula. However, the measurement of tumor size is very inaccurate. The therapeutic effects of a drug candidate can be better described as treatment-induced growth delay and specific growth delay. Another important variable in the description of tumor growth is the tumor volume doubling time. Computer programs for the calculation and description of tumor growth are also available, such as the program reported by Rygaard and Spang-Thomsen, Proc. 6th Int. Workshop on Immune-Deficient Animals, Wu and Sheng eds. (Basel, 1989), p. 301. It is noted, however, that necrosis and inflammatory responses following treatment may actually result in an increase in tumor size, at least initially. Therefore, these changes need to be carefully monitored, by a combination of a morphometric method and flow cytometric analysis.

Further, recombinant (transgenic) animal models can be engineered by introducing the coding portion of the PRO gene identified herein into the genome of animals of interest, using standard techniques for producing transgenic animals. Animals that can serve as a target for transgenic manipulation include, without limitation, mice, rats, rabbits, guinea pigs, sheep, goats, pigs, and non-human primates, *e.g.*, baboons, chimpanzees and monkeys. Techniques known in the art to introduce a transgene into such animals include pronucleic microinjection (U.S. Patent No. 4,873,191); retrovirus-mediated gene transfer into germ lines (*e.g.*, Van der Putten *et al.*, Proc. Natl. Acad. Sci. USA, **82**: 6148-615 (1985)); gene targeting in embryonic stem cells (Thompson *et al.*, Cell, **56**: 313-321 (1989)); electroporation of embryos (Lo, Mol. Cell. Biol., **3**: 1803-1814 (1983)); and sperm-mediated gene transfer. Lavitrano *et al.*, Cell, **57**: 717-73 (1989). For a review, see for example, U.S. Patent No. 4,736,866.

For the purpose of the present invention, transgenic animals include those that carry the transgene only in part of their cells ("mosaic animals"). The transgene can be integrated either as a single transgene, or in

concatamers, *e.g.*, head-to-head or head-to-tail tandems. Selective introduction of a transgene into a particular cell type is also possible by following, for example, the technique of Lasko *et al.*, Proc. Natl. Acad. Sci. USA, **89**: 6232-636 (1992). The expression of the transgene in transgenic animals can be monitored by standard techniques. For example, Southern blot analysis or PCR amplification can be used to verify the integration of the transgene. The level of mRNA expression can then be analyzed using techniques such as *in situ* hybridization, Northern blot analysis, PCR, or immunocytochemistry. The animals are further examined for signs of tumor or cancer development.

Alternatively, "knock-out" animals can be constructed that have a defective or altered gene encoding a PRO polypeptide identified herein, as a result of homologous recombination between the endogenous gene encoding the PRO polypeptide and altered genomic DNA encoding the same polypeptide introduced into an embryonic cell of the animal. For example, cDNA encoding a particular PRO polypeptide can be used to clone genomic DNA encoding that polypeptide in accordance with established techniques. A portion of the genomic DNA encoding a particular PRO polypeptide can be deleted or replaced with another gene, such as a gene encoding a selectable marker that can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector. *See, e.g.*, Thomas and Capecchi, Cell, **51**: 503 (1987) for a description of homologous recombination vectors. The vector is introduced into an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected. *See, e.g.*, Li *et al.*, Cell, **69**: 915 (1992). The selected cells are then injected into a blastocyst of an animal (*e.g.*, a mouse or rat) to form aggregation chimeras. *See, e.g.*, Bradley, in Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, E. J. Robertson, ed. (IRL: Oxford, 1987), pp. 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a "knock-out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized, for instance, by their ability to defend against certain pathological conditions and by their development of pathological conditions due to absence of the PRO polypeptide.

The efficacy of antibodies specifically binding the PRO polypeptides identified herein, and other drug candidates, can be tested also in the treatment of spontaneous animal tumors. A suitable target for such studies is the feline oral squamous cell carcinoma (SCC). Feline oral SCC is a highly invasive, malignant tumor that is the most common oral malignancy of cats, accounting for over 60% of the oral tumors reported in this species. It rarely metastasizes to distant sites, although this low incidence of metastasis may merely be a reflection of the short survival times for cats with this tumor. These tumors are usually not amenable to surgery, primarily because of the anatomy of the feline oral cavity. At present, there is no effective treatment for this tumor. Prior to entry into the study, each cat undergoes complete clinical examination and biopsy, and is scanned by computed tomography (CT). Cats diagnosed with sublingual oral squamous cell tumors are excluded from the study. The tongue can become paralyzed as a result of such tumor, and even if the treatment kills the tumor, the animals may not be able to feed themselves. Each cat is treated repeatedly, over a longer period of time. Photographs of the tumors will be taken daily during the treatment period, and at each subsequent recheck. After treatment, each cat undergoes another CT

scan. CT scans and thoracic radiograms are evaluated every 8 weeks thereafter. The data are evaluated for differences in survival, response, and toxicity as compared to control groups. Positive response may require evidence of tumor regression, preferably with improvement of quality of life and/or increased life span.

In addition, other spontaneous animal tumors, such as fibrosarcoma, adenocarcinoma, lymphoma, chondroma, or leiomyosarcoma of dogs, cats, and baboons can also be tested. Of these, mammary adenocarcinoma in dogs and cats is a preferred model as its appearance and behavior are very similar to those in humans. However, the use of this model is limited by the rare occurrence of this type of tumor in animals.

Other *in vitro* and *in vivo* cardiovascular, endothelial, and angiogenic tests known in the art are also suitable herein.

#### 5.2.4.2. Tissue Distribution

The results of the cardiovascular, endothelial, and angiogenic assays herein can be verified by further studies, such as by determining mRNA expression in various human tissues.

As noted before, gene amplification and/or gene expression in various tissues may be measured by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA (Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)), dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes.

Gene expression in various tissues, alternatively, may be measured by immunological methods, such as immunohistochemical staining of tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native-sequence PRO polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO DNA and encoding a specific antibody epitope. General techniques for generating antibodies, and special protocols for *in situ* hybridization are provided hereinbelow.

#### 5.2.4.3. Antibody Binding Studies

The results of the cardiovascular, endothelial, and angiogenic study can be further verified by antibody binding studies, in which the ability of anti-PRO antibodies to inhibit the effect of the PRO polypeptides on endothelial cells or other cells used in the cardiovascular, endothelial, and angiogenic assays is tested. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies, the preparation of which will be described hereinbelow.

Antibody binding studies may be carried out in any known assay method, such as competitive binding assays, direct and indirect sandwich assays, and immunoprecipitation assays. Zola, Monoclonal Antibodies: A Manual of Techniques (CRC Press, Inc., 1987), pp.147-158.

Competitive binding assays rely on the ability of a labeled standard to compete with the test sample analyte for binding with a limited amount of antibody. The amount of target protein in the test sample is inversely proportional to the amount of standard that becomes bound to the antibodies. To facilitate determining the amount of standard that becomes bound, the antibodies preferably are insolubilized before or after the competition, so that the standard and analyte that are bound to the antibodies may conveniently be separated from the standard and analyte that remain unbound.

Sandwich assays involve the use of two antibodies, each capable of binding to a different immunogenic portion, or epitope, of the protein to be detected. In a sandwich assay, the test sample analyte is bound by a first antibody that is immobilized on a solid support, and thereafter a second antibody binds to the analyte, thus forming an insoluble three-part complex. *See, e.g.,* U.S. Pat. No. 4,376,110. The second antibody may itself be labeled with a detectable moiety (direct sandwich assays) or may be measured using an anti-immunoglobulin antibody that is labeled with a detectable moiety (indirect sandwich assay). For example, one type of sandwich assay is an ELISA assay, in which case the detectable moiety is an enzyme.

For immunohistochemistry, the tissue sample may be fresh or frozen or may be embedded in paraffin and fixed with a preservative such as formalin, for example.

#### 5.2.4.4. Cell-Based Tumor Assays

Cell-based assays and animal models for cardiovascular, endothelial, and angiogenic disorders, such as tumors, can be used to verify the findings of a cardiovascular, endothelial, and angiogenic assay herein, and further to understand the relationship between the genes identified herein and the development and pathogenesis of undesirable cardiovascular, endothelial, and angiogenic cell growth. The role of gene products identified herein in the development and pathology of undesirable cardiovascular, endothelial, and angiogenic cell growth, *e.g.,* tumor cells, can be tested by using cells or cells lines that have been identified as being stimulated or inhibited by the PRO polypeptide herein. Such cells include, for example, those set forth in the Examples below.

In a different approach, cells of a cell type known to be involved in a particular cardiovascular, endothelial, and angiogenic disorder are transfected with the cDNAs herein, and the ability of these cDNAs to induce excessive growth or inhibit growth is analyzed. If the cardiovascular, endothelial, and angiogenic disorder is cancer, suitable tumor cells include, for example, stable tumor cell lines such as the B104-1-1 cell line (stable NIH-3T3 cell line transfected with the *neu* protooncogene) and *ras*-transfected NIH-3T3 cells, which can be transfected with the desired gene and monitored for tumorigenic growth. Such transfected cell lines can then be used to test the ability of poly- or monoclonal antibodies or antibody compositions to inhibit tumorigenic cell growth by exerting cytostatic or cytotoxic activity on the growth of the transformed cells, or by mediating antibody-dependent cellular cytotoxicity (ADCC). Cells transfected with the coding sequences of the genes identified herein can further be used to identify drug candidates for the treatment of cardiovascular, endothelial, and angiogenic disorders such as cancer.

In addition, primary cultures derived from tumors in transgenic animals (as described above) can be used in the cell-based assays herein, although stable cell lines are preferred. Techniques to derive continuous cell lines from transgenic animals are well known in the art. *See, e.g.,* Small *et al.*, Mol. Cell. Biol., **5**: 642-648 (1985).

#### 5.2.4.5. Gene Therapy

Described below are methods and compositions whereby disease symptoms may be ameliorated. Certain diseases are brought about, at least in part, by an excessive level of gene product, or by the presence of a gene product exhibiting an abnormal or excessive activity. As such, the reduction in the level and/or activity of such gene products would bring about the amelioration of such disease symptoms.

Alternatively, certain other diseases are brought about, at least in part, by the absence or reduction of the level of gene expression, or a reduction in the level of a gene product's activity. As such, an increase in the level of gene expression and/or the activity of such gene products would bring about the amelioration of such disease symptoms.

In some cases, the up-regulation of a gene in a disease state reflects a protective role for that gene product in responding to the disease condition. Enhancement of such a target gene's expression, or the activity of the target gene product, will reinforce the protective effect it exerts. Some disease states may result from an abnormally low level of activity of such a protective gene. In these cases also, an increase in the level of gene expression and/or the activity of such gene products would bring about the amelioration of such disease symptoms.

The PRO polypeptides described herein and polypeptidyl agonists and antagonists may be employed in accordance with the present invention by expression of such polypeptides *in vivo*, which is often referred to as gene therapy.

There are two major approaches to getting the nucleic acid (optionally contained in a vector) into the patient's cells: *in vivo* and *ex vivo*. For *in vivo* delivery the nucleic acid is injected directly into the patient, usually at the sites where the PRO polypeptide is required, *i.e.*, the site of synthesis of the PRO polypeptide, if known, and the site (*e.g.*, wound) where biological activity of the PRO polypeptide is needed. For *ex vivo* treatment, the patient's cells are removed, the nucleic acid is introduced into these isolated cells, and the modified cells are administered to the patient either directly or, for example, encapsulated within porous membranes that are implanted into the patient (*see, e.g.*, U.S. Pat. Nos. 4,892,538 and 5,283,187). There are a variety of techniques available for introducing nucleic acids into viable cells. The techniques vary depending upon whether the nucleic acid is transferred into cultured cells *in vitro*, or transferred *in vivo* in the cells of the intended host. Techniques suitable for the transfer of nucleic acid into mammalian cells *in vitro* include the use of liposomes, electroporation, microinjection, transduction, cell fusion, DEAE-dextran, the calcium phosphate precipitation method, etc. Transduction involves the association of a replication-defective, recombinant viral (preferably retroviral) particle with a cellular receptor, followed by introduction of the nucleic acids contained by the particle into the cell. A commonly used vector for *ex vivo* delivery of the gene is a retrovirus.

The currently preferred *in vivo* nucleic acid transfer techniques include transfection with viral or non-viral vectors (such as adenovirus, lentivirus, Herpes simplex I virus, or adeno-associated virus (AAV)) and lipid-based systems (useful lipids for lipid-mediated transfer of the gene are, for example, DOTMA, DOPE, and DC-Chol; *see, e.g.*, Tonkinson *et al.*, Cancer Investigation, 14(1): 54-65 (1996)). The most preferred vectors for use in gene therapy are viruses, most preferably adenoviruses, AAV, lentiviruses, or retroviruses. A viral vector such as a retroviral vector includes at least one transcriptional promoter/enhancer or locus-defining element(s), or other



elements that control gene expression by other means such as alternate splicing, nuclear RNA export, or post-translational modification of messenger. In addition, a viral vector such as a retroviral vector includes a nucleic acid molecule that, when transcribed in the presence of a gene encoding the PRO polypeptide, is operably linked thereto and acts as a translation initiation sequence. Such vector constructs also include a packaging signal, long terminal repeats (LTRs) or portions thereof, and positive and negative strand primer binding sites appropriate to the virus used (if these are not already present in the viral vector). In addition, such vector typically includes a signal sequence for secretion of the PRO polypeptide from a host cell in which it is placed. Preferably the signal sequence for this purpose is a mammalian signal sequence, most preferably the native signal sequence for the PRO polypeptide. Optionally, the vector construct may also include a signal that directs polyadenylation, as well as one or more restriction sites and a translation termination sequence. By way of example, such vectors will typically include a 5' LTR, a tRNA binding site, a packaging signal, an origin of second-strand DNA synthesis, and a 3' LTR or a portion thereof. Other vectors can be used that are non-viral, such as cationic lipids, polylysine, and dendrimers.

In some situations, it is desirable to provide the nucleic acid source with an agent that targets the target cells, such as an antibody specific for a cell-surface membrane protein or the target cell, a ligand for a receptor on the target cell, etc. Where liposomes are employed, proteins that bind to a cell-surface membrane protein associated with endocytosis may be used for targeting and/or to facilitate uptake, *e.g.*, capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins that undergo internalization in cycling, and proteins that target intracellular localization and enhance intracellular half-life. The technique of receptor-mediated endocytosis is described, for example, by Wu *et al.*, J. Biol. Chem., **262**: 4429-4432 (1987); and Wagner *et al.*, Proc. Natl. Acad. Sci. USA, **87**: 3410-3414 (1990). For a review of the currently known gene marking and gene therapy protocols, see, Anderson *et al.*, Science, **256**: 808-813 (1992). See also WO 93/25673 and the references cited therein.

Suitable gene therapy and methods for making retroviral particles and structural proteins can be found in, *e.g.*, U.S. Pat. No. 5,681,746.

#### 5.2.4.6. Use of Gene as a Diagnostic

This invention is also related to the use of the gene encoding the PRO polypeptide as a diagnostic. Detection of a mutated form of the PRO polypeptide will allow a diagnosis of a cardiovascular, endothelial, and angiogenic disease or a susceptibility to a cardiovascular, endothelial, and angiogenic disease, such as a tumor, since mutations in the PRO polypeptide may cause tumors.

Individuals carrying mutations in the genes encoding a human PRO polypeptide may be detected at the DNA level by a variety of techniques. Nucleic acids for diagnosis may be obtained from a patient's cells, such as from blood, urine, saliva, tissue biopsy, and autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR (Saiki *et al.*, Nature, **324**: 163-166 (1986)) prior to analysis. RNA or cDNA may also be used for the same purpose. As an example, PCR primers complementary to the nucleic acid encoding the PRO polypeptide can be used to identify and analyze the PRO polypeptide mutations. For example, deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal

genotype. Point mutations can be identified by hybridizing amplified DNA to radiolabeled RNA encoding the PRO polypeptide, or alternatively, radiolabeled antisense DNA sequences encoding the PRO polypeptide. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase A digestion or by differences in melting temperatures.

5 Genetic testing based on DNA sequence differences may be achieved by detection of alteration in electrophoretic mobility of DNA fragments in gels with or without denaturing agents. Small sequence deletions and insertions can be visualized by high resolution gel electrophoresis. DNA fragments of different sequences may be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures. See, 10 e.g., Myers *et al.*, Science, 230: 1242 (1985).

Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method, for example, Cotton *et al.*, Proc. Natl. Acad. Sci. USA, 85: 4397-4401 (1985).

15 Thus, the detection of a specific DNA sequence may be achieved by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing, or the use of restriction enzymes, e.g., restriction fragment length polymorphisms (RFLP), and Southern blotting of genomic DNA.

#### 5.2.4.7. Use to Detect PRO Polypeptide Levels

In addition to more conventional gel-electrophoresis and DNA sequencing, mutations can also be detected by *in situ* analysis.

20 Expression of nucleic acid encoding the PRO polypeptide may be linked to vascular disease or neovascularization associated with tumor formation. If the PRO polypeptide has a signal sequence and the mRNA is highly expressed in endothelial cells and to a lesser extent in smooth muscle cells, this indicates that the PRO polypeptide is present in serum. Accordingly, an anti-PRO polypeptide antibody could be used to diagnose vascular disease or neovascularization associated with tumor formation, since an altered level of this PRO polypeptide may 25 be indicative of such disorders.

A competition assay may be employed wherein antibodies specific to the PRO polypeptide are attached to a solid support and the labeled PRO polypeptide and a sample derived from the host are passed over the solid support and the amount of label detected attached to the solid support can be correlated to a quantity of the PRO polypeptide in the sample.

#### 30 5.2.4.8. Chromosome Mapping

The sequences of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Moreover, there is a current need for identifying particular sites on the chromosome. Few chromosome marking reagents based on actual sequence data (repeat polymorphisms) are presently available for marking chromosomal 35 location. The mapping of DNAs to chromosomes according to the present invention is an important first step in

correlating those sequences with genes associated with disease.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the cDNA. Computer analysis for the 3'- untranslated region is used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the primer will yield an amplified fragment.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular DNA to a particular chromosome. Using the present invention with the same oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes or pools of large genomic clones in an analogous manner. Other mapping strategies that can similarly be used to map to its chromosome include *in situ* hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome-specific cDNA libraries.

Fluorescence *in situ* hybridization (FISH) of a cDNA clone to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with cDNA as short as 500 or 600 bases; however, clones larger than 2,000 bp have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. FISH requires use of the clones from which the gene encoding the PRO polypeptide was derived, and the longer the better. For example, 2,000 bp is good, 4,000 bp is better, and more than 4,000 is probably not necessary to get good results a reasonable percentage of the time. For a review of this technique, see, Verma *et al.*, Human Chromosomes: a Manual of Basic Techniques (Pergamon Press, New York, 1988).

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available online through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region is then identified through linkage analysis (coinheritance of physically adjacent genes).

Next, it is necessary to determine the differences in the cDNA or genomic sequence between affected and unaffected individuals. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

With current resolution of physical mapping and genetic mapping techniques, a cDNA precisely localized to a chromosomal region associated with the disease could be one of between 50 and 500 potential causative genes. (This assumes 1 megabase mapping resolution and one gene per 20 kb).

#### 5.2.4.9. Screening Assays for Drug Candidates

This invention encompasses methods of screening compounds to identify those that mimic the PRO polypeptide (agonists) or prevent the effect of the PRO polypeptide (antagonists). Screening assays for antagonist drug candidates are designed to identify compounds that bind or complex with the PRO polypeptide encoded by the genes identified herein, or otherwise interfere with the interaction of the encoded polypeptides with other

cellular proteins. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates.

The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays, and cell-based assays, which are well characterized in the art.

5 All assays for antagonists are common in that they call for contacting the drug candidate with a PRO polypeptide encoded by a nucleic acid identified herein under conditions and for a time sufficient to allow these two components to interact.

10 In binding assays, the interaction is binding and the complex formed can be isolated or detected in the reaction mixture. In a particular embodiment, the PRO polypeptide encoded by the gene identified herein or the drug candidate is immobilized on a solid phase, *e.g.*, on a microtiter plate, by covalent or non-covalent attachments. Non-covalent attachment generally is accomplished by coating the solid surface with a solution of the PRO polypeptide and drying. Alternatively, an immobilized antibody, *e.g.*, a monoclonal antibody, specific for the PRO polypeptide to be immobilized can be used to anchor it to a solid surface. The assay is performed by adding the non-immobilized component, which may be labeled by a detectable label, to the immobilized component, *e.g.*, the  
15 coated surface containing the anchored component. When the reaction is complete, the non-reacted components are removed, *e.g.*, by washing, and complexes anchored on the solid surface are detected. When the originally non-immobilized component carries a detectable label, the detection of label immobilized on the surface indicates that complexing occurred. Where the originally non-immobilized component does not carry a label, complexing can be detected, for example, by using a labeled antibody specifically binding the immobilized complex.

20 If the candidate compound interacts with but does not bind to a particular PRO polypeptide encoded by a gene identified herein, its interaction with that polypeptide can be assayed by methods well known for detecting protein-protein interactions. Such assays include traditional approaches, such as, *e.g.*, cross-linking, co-immunoprecipitation, and co-purification through gradients or chromatographic columns. In addition, protein-protein interactions can be monitored by using a yeast-based genetic system described by Fields and co-workers  
25 (Fields and Song, *Nature (London)*, 340: 245-246 (1989); Chien *et al.*, *Proc. Natl. Acad. Sci. USA*, 88: 9578-9582 (1991)) as disclosed by Chevray and Nathans, *Proc. Natl. Acad. Sci. USA*, 89: 5789-5793 (1991). Many transcriptional activators, such as yeast GAL4, consist of two physically discrete modular domains, one acting as the DNA-binding domain, the other one functioning as the transcription-activation domain. The yeast expression system described in the foregoing publications (generally referred to as the "two-hybrid system") takes advantage  
30 of this property, and employs two hybrid proteins, one in which the target protein is fused to the DNA-binding domain of GAL4, and another, in which candidate activating proteins are fused to the activation domain. The expression of a GAL1-*lacZ* reporter gene under control of a GAL4-activated promoter depends on reconstitution of GAL4 activity via protein-protein interaction. Colonies containing interacting polypeptides are detected with a chromogenic substrate for  $\beta$ -galactosidase. A complete kit (MATCHMAKER™) for identifying protein-protein  
35 interactions between two specific proteins using the two-hybrid technique is commercially available from Clontech. This system can also be extended to map protein domains involved in specific protein interactions as well as to pinpoint amino acid residues that are crucial for these interactions.

Compounds that interfere with the interaction of a gene encoding a PRO polypeptide identified herein and other intra- or extracellular components can be tested as follows: usually a reaction mixture is prepared containing the product of the gene and the intra- or extracellular component under conditions and for a time allowing for the interaction and binding of the two products. To test the ability of a candidate compound to inhibit binding, the reaction is run in the absence and in the presence of the test compound. In addition, a placebo may be added to a third reaction mixture, to serve as positive control. The binding (complex formation) between the test compound and the intra- or extracellular component present in the mixture is monitored as described hereinabove. The formation of a complex in the control reaction(s) but not in the reaction mixture containing the test compound indicates that the test compound interferes with the interaction of the test compound and its reaction partner.

If the PRO polypeptide has the ability to stimulate the proliferation of endothelial cells in the presence of the co-mitogen ConA, then one example of a screening method takes advantage of this ability. Specifically, in the proliferation assay, human umbilical vein endothelial cells are obtained and cultured in 96-well flat-bottomed culture plates (Costar, Cambridge, MA) and supplemented with a reaction mixture appropriate for facilitating proliferation of the cells, the mixture containing Con-A (Calbiochem, La Jolla, CA). Con-A and the compound to be screened are added and after incubation at 37°C, cultures are pulsed with <sup>3</sup>H-thymidine and harvested onto glass fiber filters (phD; Cambridge Technology, Watertown, MA). Mean <sup>3</sup>H-thymidine incorporation (cpm) of triplicate cultures is determined using a liquid scintillation counter (Beckman Instruments, Irvine, CA). Significant <sup>3</sup>(H)-thymidine incorporation indicates stimulation of endothelial cell proliferation.

To assay for antagonists, the assay described above is performed; however, in this assay the PRO polypeptide is added along with the compound to be screened and the ability of the compound to inhibit <sup>3</sup>(H)thymidine incorporation in the presence of the PRO polypeptide indicates that the compound is an antagonist to the PRO polypeptide. Alternatively, antagonists may be detected by combining the PRO polypeptide and a potential antagonist with membrane-bound PRO polypeptide receptors or recombinant receptors under appropriate conditions for a competitive inhibition assay. The PRO polypeptide can be labeled, such as by radioactivity, such that the number of PRO polypeptide molecules bound to the receptor can be used to determine the effectiveness of the potential antagonist. The gene encoding the receptor can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Coligan *et al.*, Current Protocols in Immun., 1(2): Chapter 5 (1991). Preferably, expression cloning is employed wherein polyadenylated RNA is prepared from a cell responsive to the PRO polypeptide and a cDNA library created from this RNA is divided into pools and used to transfect COS cells or other cells that are not responsive to the PRO polypeptide. Transfected cells that are grown on glass slides are exposed to the labeled PRO polypeptide. The PRO polypeptide can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase. Following fixation and incubation, the slides are subjected to autoradiographic analysis. Positive pools are identified and sub-pools are prepared and re-transfected using an interactive sub-pooling and re-screening process, eventually yielding a single clone that encodes the putative receptor.

As an alternative approach for receptor identification, the labeled PRO polypeptide can be photoaffinity-linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is

resolved by PAGE and exposed to X-ray film. The labeled complex containing the receptor can be excised, resolved into peptide fragments, and subjected to protein micro-sequencing. The amino acid sequence obtained from micro-sequencing would be used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the gene encoding the putative receptor.

5 In another assay for antagonists, mammalian cells or a membrane preparation expressing the receptor would be incubated with the labeled PRO polypeptide in the presence of the candidate compound. The ability of the compound to enhance or block this interaction could then be measured.

The compositions useful in the treatment of cardiovascular, endothelial, and angiogenic disorders include, without limitation, antibodies, small organic and inorganic molecules, peptides, phosphopeptides, antisense and  
10 ribozyme molecules, triple-helix molecules, etc., that inhibit the expression and/or activity of the target gene product.

More specific examples of potential antagonists include an oligonucleotide that binds to the fusions of immunoglobulin with a PRO polypeptide, and, in particular, antibodies including, without limitation, poly- and monoclonal antibodies and antibody fragments, single-chain antibodies, anti-idiotypic antibodies, and chimeric or  
15 humanized versions of such antibodies or fragments, as well as human antibodies and antibody fragments. Alternatively, a potential antagonist may be a closely related protein, for example, a mutated form of the PRO polypeptide that recognizes the receptor but imparts no effect, thereby competitively inhibiting the action of the PRO polypeptide.

Another potential PRO polypeptide antagonist is an antisense RNA or DNA construct prepared using  
20 antisense technology, where, *e.g.*, an antisense RNA or DNA molecule acts to block directly the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes the mature PRO polypeptides herein, is used to design an antisense RNA oligonucleotide  
25 of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple helix - *see, Lee et al., Nucl. Acids Res.*, 6:3073 (1979); Cooney *et al., Science*, 241: 456 (1988); Dervan *et al., Science*, 251:1360 (1991)), thereby preventing transcription and the production of the PRO polypeptide. A sequence "complementary" to a portion of an RNA, as referred to herein, means a sequence having sufficient complementarity to be able to hybridize with the RNA, forming a stable duplex;  
30 in the case of double-stranded antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex helix formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point  
35 of the hybridized complex. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into the PRO polypeptide (antisense - Okano, Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression (CRC Press: Boca Raton, FL, 1988).

The antisense oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*),  
 5 or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger, *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre, *et al.*, 1987, *Proc. Natl. Acad. Sci. U.S.A.* 84:648-652; PCT Publication No. WO88/09810, published December 15, 1988) or the blood-brain barrier (see, *e.g.*, PCT Publication No. WO89/10134, published April 25, 1988), hybridization-triggered cleavage agents (see, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents (see, *e.g.*, Zon, 1988, *Pharm. Res.* 5:539-549). To this end, the  
 10 oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine,  
 15 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil,  
 20 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate,  
 25 a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet another embodiment, the antisense oligonucleotide is an  $\alpha$ -anomeric oligonucleotide. An  $\alpha$ -anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each other (Gautier, *et al.*, 1987, *Nucl. Acids Res.* 15:6625-6641). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue, *et al.*, 1987, *Nucl. Acids Res.* 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue, *et al.*, 1987, *FEBS Lett.* 215:327-330).  
 30

Oligonucleotides of the invention may be synthesized by standard methods known in the art, *e.g.*, by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.).  
 35 As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein, *et al.* (1988, *Nucl. Acids Res.* 16:3209), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin, *et al.*, 1988, *Proc. Natl. Acad. Sci. U.S.A.* 85:7448-7451), etc.

The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of the PRO polypeptide. When antisense DNA is used, oligodeoxyribonucleotides derived from the translation-initiation site, *e.g.*, between about -10 and +10 positions of the target gene nucleotide sequence, are preferred.

5 Antisense RNA or DNA molecules are generally at least about 5 bases in length, about 10 bases in length, about 15 bases in length, about 20 bases in length, about 25 bases in length, about 30 bases in length, about 35 bases in length, about 40 bases in length, about 45 bases in length, about 50 bases in length, about 55 bases in length, about 60 bases in length, about 65 bases in length, about 70 bases in length, about 75 bases in length, about 80 bases in length, about 85 bases in length, about 90 bases in length, about 95 bases in length, about 100 bases in length, or more.

10 Potential antagonists further include small molecules that bind to the active site, the receptor binding site, or growth factor or other relevant binding site of the PRO polypeptide, thereby blocking the normal biological activity of the PRO polypeptide. Examples of small molecules include, but are not limited to, small peptides or peptide-like molecules, preferably soluble peptides, and synthetic non-peptidyl organic or inorganic compounds.

15 Additional potential antagonists are ribozymes, which are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. Ribozymes act by sequence-specific hybridization to the complementary target RNA, followed by endonucleolytic cleavage. Specific ribozyme cleavage sites within a potential RNA target can be identified by known techniques. For further details *see, e.g.*, Rossi, Current Biology, 4: 469-471 (1994), and PCT publication No. WO 97/33551 (published September 18, 1997).

20 While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy target gene mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions which form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Myers, 1995, *Molecular Biology and Biotechnology: A Comprehensive Desk Reference*, VCH Publishers, New York, (see especially Figure 4, page 833) and in Haseloff and Gerlach, 1988, *Nature*, 334:585-591, which is incorporated herein by reference in its entirety.

25 Preferably the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the target gene mRNA, *i.e.*, to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

30 The ribozymes of the present invention also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one which occurs naturally in *Tetrahymena thermophila* (known as the IVS, or L-19 IVS RNA) and which has been extensively described by Thomas Cech and collaborators (Zaug, *et al.*, 1984, *Science*, 224:574-578; Zaug and Cech, 1986, *Science*, 231:470-475; Zaug, *et al.*, 1986, *Nature*, 324:429-433; published International patent application No. WO 88/04300 by University Patents Inc.; Been and Cech, 1986, *Cell*, 47:207-216). The Cech-type ribozymes have an eight base pair active site that hybridizes to a target RNA sequence



whereafter cleavage of the target RNA takes place. The invention encompasses those Cech-type ribozymes that target eight base-pair active site sequences that are present in the target gene.

As in the antisense approach, the ribozymes can be composed of modified oligonucleotides (*e.g.*, for improved stability, targeting, *etc.*) and should be delivered to cells that express the target gene *in vivo*. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy endogenous target gene messages and inhibit translation. Because ribozymes, unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

Nucleic acid molecules in triple-helix formation used to inhibit transcription should be single-stranded and composed of deoxynucleotides. The base composition of these oligonucleotides is designed such that it promotes triple-helix formation via Hoogsteen base-pairing rules, which generally require sizeable stretches of purines or pyrimidines on one strand of a duplex. For further details *see, e.g.*, PCT publication No. WO 97/33551, *supra*.

These small molecules can be identified by any one or more of the screening assays discussed hereinabove and/or by any other screening techniques well known for those skilled in the art.

#### 5.2.4.10. Types of Cardiovascular, Endothelial, and Angiogenic Disorders to be Treated

The PRO polypeptides, or agonists or antagonists thereto, that have activity in the cardiovascular, angiogenic, and endothelial assays described herein, and/or whose gene product has been found to be localized to the cardiovascular system, are likely to have therapeutic uses in a variety of cardiovascular, endothelial, and angiogenic disorders, including systemic disorders that affect vessels, such as diabetes mellitus. Their therapeutic utility could include diseases of the arteries, capillaries, veins, and/or lymphatics. Examples of treatments hereunder include treating muscle wasting disease, treating osteoporosis, aiding in implant fixation to stimulate the growth of cells around the implant and therefore facilitate its attachment to its intended site, increasing IGF stability in tissues or in serum, if applicable, and increasing binding to the IGF receptor (since IGF has been shown *in vitro* to enhance human marrow erythroid and granulocytic progenitor cell growth).

The PRO polypeptides or agonists or antagonists thereto may also be employed to stimulate erythropoiesis or granulopoiesis, to stimulate wound healing or tissue regeneration and associated therapies concerned with re-growth of tissue, such as connective tissue, skin, bone, cartilage, muscle, lung, or kidney, to promote angiogenesis, to stimulate or inhibit migration of endothelial cells, and to proliferate the growth of vascular smooth muscle and endothelial cell production. The increase in angiogenesis mediated by the PRO polypeptide or agonist would be beneficial to ischemic tissues and to collateral coronary development in the heart subsequent to coronary stenosis. Antagonists are used to inhibit the action of such polypeptides, for example, to limit the production of excess connective tissue during wound healing or pulmonary fibrosis if the PRO polypeptide promotes such production. This would include treatment of acute myocardial infarction and heart failure.

Moreover, the present invention provides the treatment of cardiac hypertrophy, regardless of the underlying cause, by administering a therapeutically effective dose of the PRO polypeptide, or agonist or antagonist thereto. If the objective is the treatment of human patients, the PRO polypeptide preferably is recombinant human PRO

polypeptide (rhPRO polypeptide). The treatment for cardiac hypertrophy can be performed at any of its various stages, which may result from a variety of diverse pathologic conditions, including myocardial infarction, hypertension, hypertrophic cardiomyopathy, and valvular regurgitation. The treatment extends to all stages of the progression of cardiac hypertrophy, with or without structural damage of the heart muscle, regardless of the underlying cardiac disorder.

The decision of whether to use the molecule itself or an agonist thereof for any particular indication, as opposed to an antagonist to the molecule, would depend mainly on whether the molecule herein promotes cardiovascularization, genesis of endothelial cells, or angiogenesis or inhibits these conditions. For example, if the molecule promotes angiogenesis, an antagonist thereof would be useful for treatment of disorders where it is desired to limit or prevent angiogenesis. Examples of such disorders include vascular tumors such as haemangioma, tumor angiogenesis, neovascularization in the retina, choroid, or cornea, associated with diabetic retinopathy or premature infant retinopathy or macular degeneration and proliferative vitreoretinopathy, rheumatoid arthritis, Crohn's disease, atherosclerosis, ovarian hyperstimulation, psoriasis, endometriosis associated with neovascularization, restenosis subsequent to balloon angioplasty, scar tissue overproduction, for example, that seen in a keloid that forms after surgery, fibrosis after myocardial infarction, or fibrotic lesions associated with pulmonary fibrosis.

If, however, the molecule inhibits angiogenesis, it would be expected to be used directly for treatment of the above conditions.

On the other hand, if the molecule stimulates angiogenesis it would be used itself (or an agonist thereof) for indications where angiogenesis is desired such as peripheral vascular disease, hypertension, inflammatory vasculitides, Reynaud's disease and Reynaud's phenomenon, aneurysms, arterial restenosis, thrombophlebitis, lymphangitis, lymphedema, wound healing and tissue repair, ischemia reperfusion injury, angina, myocardial infarctions such as acute myocardial infarctions, chronic heart conditions, heart failure such as congestive heart failure, and osteoporosis.

If, however, the molecule inhibits angiogenesis, an antagonist thereof would be used for treatment of those conditions where angiogenesis is desired.

Specific types of diseases are described below, where the PRO polypeptide herein or agonists or antagonists thereof may serve as useful for vascular-related drug targeting or as therapeutic targets for the treatment or prevention of the disorders. Atherosclerosis is a disease characterized by accumulation of plaques of intimal thickening in arteries, due to accumulation of lipids, proliferation of smooth muscle cells, and formation of fibrous tissue within the arterial wall. The disease can affect large, medium, and small arteries in any organ. Changes in endothelial and vascular smooth muscle cell function are known to play an important role in modulating the accumulation and regression of these plaques.

Hypertension is characterized by raised vascular pressure in the systemic arterial, pulmonary arterial, or portal venous systems. Elevated pressure may result from or result in impaired endothelial function and/or vascular disease.

Inflammatory vasculitides include giant cell arteritis, Takayasu's arteritis, polyarteritis nodosa (including the microangiopathic form), Kawasaki's disease, microscopic polyangiitis, Wegener's granulomatosis, and a variety

of infectious-related vascular disorders (including Henoch-Schonlein purpura). Altered endothelial cell function has been shown to be important in these diseases.

Reynaud's disease and Reynaud's phenomenon are characterized by intermittent abnormal impairment of the circulation through the extremities on exposure to cold. Altered endothelial cell function has been shown to be important in this disease.

Aneurysms are saccular or fusiform dilatations of the arterial or venous tree that are associated with altered endothelial cell and/or vascular smooth muscle cells.

Arterial restenosis (restenosis of the arterial wall) may occur following angioplasty as a result of alteration in the function and proliferation of endothelial and vascular smooth muscle cells.

Thrombophlebitis and lymphangitis are inflammatory disorders of veins and lymphatics, respectively, that may result from, and/or in, altered endothelial cell function. Similarly, lymphedema is a condition involving impaired lymphatic vessels resulting from endothelial cell function.

The family of benign and malignant vascular tumors are characterized by abnormal proliferation and growth of cellular elements of the vascular system. For example, lymphangiomas are benign tumors of the lymphatic system that are congenital, often cystic, malformations of the lymphatics that usually occur in newborns. Cystic tumors tend to grow into the adjacent tissue. Cystic tumors usually occur in the cervical and axillary region. They can also occur in the soft tissue of the extremities. The main symptoms are dilated, sometimes reticular, structured lymphatics and lymphocysts surrounded by connective tissue. Lymphangiomas are assumed to be caused by improperly connected embryonic lymphatics or their deficiency. The result is impaired local lymph drainage. Griener *et al.*, *Lymphology*, 4: 140-144 (1971).

Another use for the PRO polypeptides herein or agonists or antagonists thereto is in the prevention of tumor angiogenesis, which involves vascularization of a tumor to enable it to grow and/or metastasize. This process is dependent on the growth of new blood vessels. Examples of neoplasms and related conditions that involve tumor angiogenesis include breast carcinomas, lung carcinomas, gastric carcinomas, esophageal carcinomas, colorectal carcinomas, liver carcinomas, ovarian carcinomas, thecomas, arrhenoblastomas, cervical carcinomas, endometrial carcinoma, endometrial hyperplasia, endometriosis, fibrosarcomas, choriocarcinoma, head and neck cancer, nasopharyngeal carcinoma, laryngeal carcinomas, hepatoblastoma, Kaposi's sarcoma, melanoma, skin carcinomas, hemangioma, cavernous hemangioma, hemangioblastoma, pancreas carcinomas, retinoblastoma, astrocytoma, glioblastoma, Schwannoma, oligodendroglioma, medulloblastoma, neuroblastomas, rhabdomyosarcoma, osteogenic sarcoma, leiomyosarcomas, urinary tract carcinomas, thyroid carcinomas, Wilm's tumor, renal cell carcinoma, prostate carcinoma, abnormal vascular proliferation associated with phakomatoses, edema (such as that associated with brain tumors), and Meigs' syndrome.

Age-related macular degeneration (AMD) is a leading cause of severe visual loss in the elderly population. The exudative form of AMD is characterized by choroidal neovascularization and retinal pigment epithelial cell detachment. Because choroidal neovascularization is associated with a dramatic worsening in prognosis, the PRO polypeptide or agonist or antagonist thereto is expected to be useful in reducing the severity of AMD.

Healing of trauma such as wound healing and tissue repair is also a targeted use for the PRO polypeptides herein or their agonists or antagonists. Formation and regression of new blood vessels is essential for tissue healing and repair. This category includes bone, cartilage, tendon, ligament, and/or nerve tissue growth or regeneration, as well as wound healing and tissue repair and replacement, and in the treatment of burns, incisions, and ulcers.

5 A PRO polypeptide or agonist or antagonist thereof that induces cartilage and/or bone growth in circumstances where bone is not normally formed has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a PRO polypeptide or agonist or antagonist thereof may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma-

10 induced, or oncologic, resection-induced craniofacial defects, and also is useful in cosmetic plastic surgery. PRO polypeptides or agonists or antagonists thereto may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

15 It is expected that a PRO polypeptide or agonist or antagonist thereto may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, or endothelium), muscle (smooth, skeletal, or cardiac), and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate.

20 A PRO polypeptide herein or agonist or antagonist thereto may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage. Also, the PRO polypeptide or agonist or antagonist thereto may be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells, or for inhibiting the growth of tissues described above.

25 A PRO polypeptide or agonist or antagonist thereto may also be used in the treatment of periodontal diseases and in other tooth-repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells, or induce differentiation of progenitors of bone-forming cells. A PRO polypeptide herein or an agonist or an antagonist thereto may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes, since blood

30 vessels play an important role in the regulation of bone turnover and growth. Another category of tissue regeneration activity that may be attributable to the PRO polypeptide herein or agonist or antagonist thereto is tendon/ligament formation. A protein that induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed has application in the healing of tendon or ligament tears, deformities, and other tendon or ligament defects in humans and other animals. Such

35 a preparation may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. *De novo* tendon/ligament-like tissue formation induced by a composition of the PRO polypeptide herein or

agonist or antagonist thereto contributes to the repair of congenital, trauma-induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions herein may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions herein may also be useful in the treatment of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The PRO polypeptide or its agonist or antagonist may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.*, for the treatment of central and peripheral nervous system disease and neuropathies, as well as mechanical and traumatic disorders, that involve degeneration, death, or trauma to neural cells or nerve tissue. More specifically, a PRO polypeptide or its agonist or antagonist may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions that may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma, and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a PRO polypeptide herein or agonist or antagonist thereto.

Ischemia-reperfusion injury is another indication. Endothelial cell dysfunction may be important in both the initiation of, and in regulation of the sequelae of events that occur following ischemia-reperfusion injury.

Rheumatoid arthritis is a further indication. Blood vessel growth and targeting of inflammatory cells through the vasculature is an important component in the pathogenesis of rheumatoid and sero-negative forms of arthritis.

A PRO polypeptide or its agonist or antagonist may also be administered prophylactically to patients with cardiac hypertrophy, to prevent the progression of the condition, and avoid sudden death, including death of asymptomatic patients. Such preventative therapy is particularly warranted in the case of patients diagnosed with massive left ventricular cardiac hypertrophy (a maximal wall thickness of 35 mm or more in adults, or a comparable value in children), or in instances when the hemodynamic burden on the heart is particularly strong.

A PRO polypeptide or its agonist or antagonist may also be useful in the management of atrial fibrillation, which develops in a substantial portion of patients diagnosed with hypertrophic cardiomyopathy.

Further indications include angina, myocardial infarctions such as acute myocardial infarctions, and heart failure such as congestive heart failure. Additional non-neoplastic conditions include psoriasis, diabetic and other proliferative retinopathies including retinopathy of prematurity, retrolental fibroplasia, neovascular glaucoma, thyroid hyperplasias (including Grave's disease), corneal and other tissue transplantation, chronic inflammation, lung inflammation, nephrotic syndrome, preeclampsia, ascites, pericardial effusion (such as that associated with pericarditis), and pleural effusion.

In view of the above, the PRO polypeptides or agonists or antagonists thereof described herein, which are shown to alter or impact endothelial cell function, proliferation, and/or form, are likely to play an important role in the etiology and pathogenesis of many or all of the disorders noted above, and as such can serve as therapeutic targets to augment or inhibit these processes or for vascular-related drug targeting in these disorders.

5

#### 5.2.4.11. Administration Protocols, Schedules, Doses, and Formulations

The molecules herein and agonists and antagonists thereto are pharmaceutically useful as a prophylactic and therapeutic agent for various disorders and diseases as set forth above.

Therapeutic compositions of the PRO polypeptides or agonists or antagonists are prepared for storage by mixing the desired molecule having the appropriate degree of purity with optional pharmaceutically acceptable carriers, excipients, or stabilizers (Remington's Pharmaceutical Sciences, 16th edition, Osol, A. ed. (1980)), in the form of lyophilized formulations or aqueous solutions. Acceptable carriers, excipients, or stabilizers are nontoxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid and methionine; preservatives (such as octadecyldimethylbenzyl ammonium chloride; hexamethonium chloride; benzalkonium chloride, benzethonium chloride; phenol, butyl or benzyl alcohol; alkyl parabens such as methyl or propyl paraben; catechol; resorcinol; cyclohexanol; 3-pentanol; and m-cresol); low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, histidine, arginine, or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrans; chelating agents such as EDTA; sugars such as sucrose, mannitol, trehalose or sorbitol; salt-forming counter-ions such as sodium; metal complexes (*e.g.*, Zn-protein complexes); and/or non-ionic surfactants such as TWEEN™, PLURONICS™ or polyethylene glycol (PEG).

Additional examples of such carriers include ion exchangers, alumina, aluminum stearate, lecithin, serum proteins, such as human serum albumin, buffer substances such as phosphates, glycine, sorbic acid, potassium sorbate, partial glyceride mixtures of saturated vegetable fatty acids, water, salts, or electrolytes such as protamine sulfate, disodium hydrogen phosphate, potassium hydrogen phosphate, sodium chloride, zinc salts, colloidal silica, magnesium trisilicate, polyvinyl pyrrolidone, cellulose-based substances, and polyethylene glycol. Carriers for topical or gel-based forms of agonist or antagonist include polysaccharides such as sodium carboxymethylcellulose or methylcellulose, polyvinylpyrrolidone, polyacrylates, polyoxyethylene-polyoxypropylene-block polymers, polyethylene glycol, and wood wax alcohols. For all administrations, conventional depot forms are suitably used. Such forms include, for example, microcapsules, nano-capsules, liposomes, plasters, inhalation forms, nose sprays, sublingual tablets, and sustained-release preparations. The PRO polypeptides or agonists or antagonists will typically be formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml.

Another formulation comprises incorporating a PRO polypeptide or agonist or antagonist thereof into formed articles. Such articles can be used in modulating endothelial cell growth and angiogenesis. In addition, tumor invasion and metastasis may be modulated with these articles.

PRO polypeptides or agonists or antagonists to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes, prior to or following lyophilization and reconstitution. PRO polypeptides ordinarily will be stored in lyophilized form or in solution if administered systemically. If in lyophilized form, the PRO polypeptide or agonist or antagonist thereto is typically formulated in combination with other ingredients for reconstitution with an appropriate diluent at the time for use. An example of a liquid formulation of a PRO polypeptide or agonist or antagonist is a sterile, clear, colorless unpreserved solution filled in a single-dose vial for subcutaneous injection. Preserved pharmaceutical compositions suitable for repeated use may contain, for example, depending mainly on the indication and type of polypeptide:

- a) PRO polypeptide or agonist or antagonist thereto;
- b) a buffer capable of maintaining the pH in a range of maximum stability of the polypeptide or other molecule in solution, preferably about 4-8;
- c) a detergent/surfactant primarily to stabilize the polypeptide or molecule against agitation-induced aggregation;
- d) an isotonicifier;
- e) a preservative selected from the group of phenol, benzyl alcohol and a benzethonium halide, *e.g.*, chloride; and
- f) water.

If the detergent employed is non-ionic, it may, for example, be polysorbates (*e.g.*, POLYSORBATE™ (TWEEN™) 20, 80, etc.) or poloxamers (*e.g.*, POLOXAMER™ 188). The use of non-ionic surfactants permits the formulation to be exposed to shear surface stresses without causing denaturation of the polypeptide. Further, such surfactant-containing formulations may be employed in aerosol devices such as those used in a pulmonary dosing, and needleless jet injector guns (*see, e.g.*, EP 257,956).

An isotonicifier may be present to ensure isotonicity of a liquid composition of the PRO polypeptide or agonist or antagonist thereto, and includes polyhydric sugar alcohols, preferably trihydric or higher sugar alcohols, such as glycerin, erythritol, arabitol, xylitol, sorbitol, and mannitol. These sugar alcohols can be used alone or in combination. Alternatively, sodium chloride or other appropriate inorganic salts may be used to render the solutions isotonic.

The buffer may, for example, be an acetate, citrate, succinate, or phosphate buffer depending on the pH desired. The pH of one type of liquid formulation of this invention is buffered in the range of about 4 to 8, preferably about physiological pH.

The preservatives phenol, benzyl alcohol and benzethonium halides, *e.g.*, chloride, are known antimicrobial agents that may be employed.

Therapeutic PRO polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle. The formulations are preferably administered as repeated intravenous (i.v.), subcutaneous (s.c.), or intramuscular (i.m.) injections, or as aerosol formulations suitable for intranasal or intrapulmonary delivery (for intrapulmonary delivery *see, e.g.*, EP 257,956).

PRO polypeptides can also be administered in the form of sustained-released preparations. Suitable examples of sustained-release preparations include semipermeable matrices of solid hydrophobic polymers containing the protein, which matrices are in the form of shaped articles, *e.g.*, films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (*e.g.*, poly(2-hydroxyethyl-methacrylate) as described by Langer *et al.*, J. Biomed. Mater. Res., **15**: 167-277 (1981) and Langer, Chem. Tech., **12**: 98-105 (1982) or poly(vinylalcohol)), polylactides (U.S. Patent No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma ethyl-L-glutamate (Sidman *et al.*, Biopolymers, **22**: 547-556 (1983)), non-degradable ethylene-vinyl acetate (Langer *et al.*, *supra*), degradable lactic acid-glycolic acid copolymers such as the Lupron Depot™ (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(-)-3-hydroxybutyric acid (EP 133,988).

While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods. When encapsulated proteins remain in the body for a long time, they may denature or aggregate as a result of exposure to moisture at 37°C, resulting in a loss of biological activity and possible changes in immunogenicity. Rational strategies can be devised for protein stabilization depending on the mechanism involved. For example, if the aggregation mechanism is discovered to be intermolecular S-S bond formation through thio-disulfide interchange, stabilization may be achieved by modifying sulfhydryl residues, lyophilizing from acidic solutions, controlling moisture content, using appropriate additives, and developing specific polymer matrix compositions.

Sustained-release PRO polypeptide compositions also include liposomally entrapped PRO polypeptides. Liposomes containing the PRO polypeptide are prepared by methods known *per se*: DE 3,218,121; Epstein *et al.*, Proc. Natl. Acad. Sci. USA, **82**: 3688-3692 (1985); Hwang *et al.*, Proc. Natl. Acad. Sci. USA, **77**: 4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese patent application 83-118008; U.S. Patent Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. % cholesterol, the selected proportion being adjusted for the optimal therapy.

The therapeutically effective dose of a PRO polypeptide or agonist or antagonist thereto will, of course, vary depending on such factors as the pathological condition to be treated (including prevention), the method of administration, the type of compound being used for treatment, any co-therapy involved, the patient's age, weight, general medical condition, medical history, etc., and its determination is well within the skill of a practicing physician. Accordingly, it will be necessary for the therapist to titer the dosage and modify the route of administration as required to obtain the maximal therapeutic effect. If the PRO polypeptide has a narrow host range, for the treatment of human patients formulations comprising human PRO polypeptide, more preferably native-sequence human PRO polypeptide, are preferred. The clinician will administer the PRO polypeptide until a dosage is reached that achieves the desired effect for treatment of the condition in question. For example, if the objective is the treatment of CHF, the amount would be one that inhibits the progressive cardiac hypertrophy associated with this condition. The progress of this therapy is easily monitored by echo cardiography. Similarly, in patients with hypertrophic cardiomyopathy, the PRO polypeptide can be administered on an empirical basis.



With the above guidelines, the effective dose generally is within the range of from about 0.001 to about 1.0 mg/kg, more preferably about 0.01-1.0 mg/kg, most preferably about 0.01-0.1 mg/kg.

For non-oral use in treating human adult hypertension, it is advantageous to administer the PRO polypeptide in the form of an injection at about 0.01 to 50 mg, preferably about 0.05 to 20 mg, most preferably 1 to 20 mg, per kg body weight, 1 to 3 times daily by intravenous injection. For oral administration, a molecule based on the PRO polypeptide is preferably administered at about 5 mg to 1 g, preferably about 10 to 100 mg, per kg body weight, 1 to 3 times daily. It should be appreciated that endotoxin contamination should be kept minimally at a safe level, for example, less than 0.5 ng/mg protein. Moreover, for human administration, the formulations preferably meet sterility, pyrogenicity, general safety, and purity as required by FDA Office and Biologics standards.

The dosage regimen of a pharmaceutical composition containing the PRO polypeptide to be used in tissue regeneration will be determined by the attending physician considering various factors that modify the action of the polypeptides, *e.g.*, amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (*e.g.*, bone), the patient's age, sex, and diet, the severity of any infection, time of administration, and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF-I, to the final composition may also affect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations, and tetracycline labeling.

The route of PRO polypeptide or antagonist or agonist administration is in accord with known methods, *e.g.*, by injection or infusion by intravenous, intramuscular, intracerebral, intraperitoneal, intracerebrospinal, subcutaneous, intraocular, intraarticular, intrasynovial, intrathecal, oral, topical, or inhalation routes, or by sustained-release systems as noted below. The PRO polypeptide or agonist or antagonists thereof also are suitably administered by intratumoral, peritumoral, intralesional, or perilesional routes, to exert local as well as systemic therapeutic effects. The intraperitoneal route is expected to be particularly useful, for example, in the treatment of ovarian tumors.

If a peptide or small molecule is employed as an antagonist or agonist, it is preferably administered orally or non-orally in the form of a liquid or solid to mammals.

Examples of pharmacologically acceptable salts of molecules that form salts and are useful hereunder include alkali metal salts (*e.g.*, sodium salt, potassium salt), alkaline earth metal salts (*e.g.*, calcium salt, magnesium salt), ammonium salts, organic base salts (*e.g.*, pyridine salt, triethylamine salt), inorganic acid salts (*e.g.*, hydrochloride, sulfate, nitrate), and salts of organic acid (*e.g.*, acetate, oxalate, p-toluenesulfonate).

For compositions herein that are useful for bone, cartilage, tendon, or ligament regeneration, the therapeutic method includes administering the composition topically, systemically, or locally as an implant or device. When administered, the therapeutic composition for use is in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage, or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Preferably, for bone and/or cartilage formation, the composition would include a matrix capable of delivering the

protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and preferably capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

5 The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance, and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid, and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are  
10 nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above-mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

15 One specific embodiment is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the polypeptide compositions from disassociating from the matrix.

20 One suitable family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropylmethylcellulose, and carboxymethylcellulose, one preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer, and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt%, based on total formulation weight, which  
25 represents the amount necessary to prevent desorption of the polypeptide (or its antagonist) from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the polypeptide (or its antagonist) the opportunity to assist the osteogenic activity of the progenitor cells.

#### 5.2.4.12. Combination Therapies

30 The effectiveness of the PRO polypeptide or an agonist or antagonist thereof in preventing or treating the disorder in question may be improved by administering the active agent serially or in combination with another agent that is effective for those purposes, either in the same composition or as separate compositions.

For example, for treatment of cardiac hypertrophy, PRO polypeptide therapy can be combined with the administration of inhibitors of known cardiac myocyte hypertrophy factors, *e.g.*, inhibitors of  $\alpha$ -adrenergic agonists  
35 such as phenylephrine; endothelin-1 inhibitors such as BOSENTAN<sup>TM</sup> and MOXONODIN<sup>TM</sup>; inhibitors to CT-1

(U.S. Pat. No. 5,679,545); inhibitors to LIF; ACE inhibitors; des-aspartate-angiotensin I inhibitors (U.S. Pat. No. 5,773,415), and angiotensin II inhibitors.

5 For treatment of cardiac hypertrophy associated with hypertension, the PRO polypeptide can be administered in combination with  $\beta$ -adrenergic receptor blocking agents, *e.g.*, propranolol, timolol, tertalolol, carteolol, nadolol, betaxolol, penbutolol, acetobutolol, atenolol, metoprolol, or carvedilol; ACE inhibitors, *e.g.*,  
10 quinapril, captopril, enalapril, ramipril, benazepril, fosinopril, or lisinopril; diuretics, *e.g.*, chlorothiazide, hydrochlorothiazide, hydroflumethazide, methylchlorothiazide, benzthiazide, dichlorphenamide, acetazolamide, or indapamide; and/or calcium channel blockers, *e.g.*, diltiazem, nifedipine, verapamil, or nocardipine. Pharmaceutical compositions comprising the therapeutic agents identified herein by their generic names are commercially available,  
15 and are to be administered following the manufacturers' instructions for dosage, administration, adverse effects, contraindications, etc. *See, e.g., Physicians' Desk Reference* (Medical Economics Data Production Co.: Montvale, N.J., 1997), 51th Edition.

Preferred candidates for combination therapy in the treatment of hypertrophic cardiomyopathy are  $\beta$ -adrenergic-blocking drugs (*e.g.*, propranolol, timolol, tertalolol, carteolol, nadolol, betaxolol, penbutolol,  
20 acetobutolol, atenolol, metoprolol, or carvedilol), verapamil, difedipine, or diltiazem. Treatment of hypertrophy associated with high blood pressure may require the use of antihypertensive drug therapy, using calcium channel blockers, *e.g.*, diltiazem, nifedipine, verapamil, or nocardipine;  $\beta$ -adrenergic blocking agents; diuretics, *e.g.*, chlorothiazide, hydrochlorothiazide, hydroflumethazide, methylchlorothiazide, benzthiazide, dichlorphenamide, acetazolamide, or indapamide; and/or ACE-inhibitors, *e.g.*, quinapril, captopril, enalapril, ramipril, benazepril, fosinopril, or lisinopril.

For other indications, PRO polypeptides or their agonists or antagonists may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as EGF, PDGF, TGF- $\alpha$  or TGF- $\beta$ , IGF, FGF, and CTGF.

25 In addition, PRO polypeptides or their agonists or antagonists used to treat cancer may be combined with cytotoxic, chemotherapeutic, or growth-inhibitory agents as identified above. Also, for cancer treatment, the PRO polypeptide or agonist or antagonist thereof is suitably administered serially or in combination with radiological treatments, whether involving irradiation or administration of radioactive substances.

The effective amounts of the therapeutic agents administered in combination with the PRO polypeptide or agonist or antagonist thereof will be at the physician's or veterinarian's discretion. Dosage administration and  
30 adjustment is done to achieve maximal management of the conditions to be treated. For example, for treating hypertension, these amounts ideally take into account use of diuretics or digitalis, and conditions such as hyper- or hypotension, renal impairment, etc. The dose will additionally depend on such factors as the type of the therapeutic agent to be used and the specific patient being treated. Typically, the amount employed will be the same dose as that used, if the given therapeutic agent is administered without the PRO polypeptide.

#### 5.2.4.13. Articles of Manufacture

An article of manufacture such as a kit containing the PRO polypeptide or agonists or antagonists thereof useful for the diagnosis or treatment of the disorders described above comprises at least a container and a label. Suitable containers include, for example, bottles, vials, syringes, and test tubes. The containers may be formed from a variety of materials such as glass or plastic. The container holds a composition that is effective for diagnosing or treating the condition and may have a sterile access port (for example, the container may be an intravenous solution bag or a vial having a stopper pierceable by a hypodermic injection needle). The active agent in the composition is the PRO polypeptide or an agonist or antagonist thereto. The label on, or associated with, the container indicates that the composition is used for diagnosing or treating the condition of choice. The article of manufacture may further comprise a second container comprising a pharmaceutically-acceptable buffer, such as phosphate-buffered saline, Ringer's solution, and dextrose solution. It may further include other materials desirable from a commercial and user standpoint, including other buffers, diluents, filters, needles, syringes, and package inserts with instructions for use. The article of manufacture may also comprise a second or third container with another active agent as described above.

#### 5.2.5. Antibodies

Some of the most promising drug candidates according to the present invention are antibodies and antibody fragments that may inhibit the production or the gene product of the genes identified herein and/or reduce the activity of the gene products.

##### 5.2.5.1. Polyclonal Antibodies

Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the PRO polypeptide or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include, but are not limited to, keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants that may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A or synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

##### 5.2.5.2. Monoclonal Antibodies

The anti-PRO antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, *Nature*, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal is typically immunized with an

immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.

5 The immunizing agent will typically include the PRO polypeptide or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell. Goding, Monoclonal Antibodies: Principles and Practice (New York: Academic Press, 1986), pp. 59-103. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine, and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, 10 immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells:

15 Preferred immortalized cell lines are those that fuse efficiently, support stable high-level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies. Kozbor, J. Immunol., 133:3001 (1984); Brodeur *et al.*, Monoclonal Antibody Production Techniques and Applications (Marcel Dekker, Inc.: New York, 1987) pp. 51-63. 20

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the PRO polypeptide. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980). 25

After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods. Goding, *supra*. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal. 30

The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

35 The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the

invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy- and light-chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison *et al.*, *supra*) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy-chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

*In vitro* methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly Fab fragments, can be accomplished using routine techniques known in the art.

#### 5.2.5.3. Human and Humanized Antibodies

The anti-PRO antibodies may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (*e.g.*, murine) antibodies are chimeric immunoglobulins, immunoglobulin chains, or fragments thereof (such as Fv, Fab, Fab', F(ab')<sub>2</sub> or other antigen-binding subsequences of antibodies) that contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a CDR of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat, or rabbit having the desired specificity, affinity, and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues that are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin, and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody preferably also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin. Jones *et al.*, Nature, 321: 522-525 (1986); Riechmann *et al.*, Nature, 332: 323-329 (1988); Presta, Curr. Op. Struct. Biol., 2:593-596 (1992).

Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source that is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers (Jones *et al.*, Nature, 321: 522-525 (1986); Riechmann *et al.*, Nature, 332: 323-327 (1988); Verhoeyen *et al.*, Science, 239: 1534-

1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries. Hoogenboom and Winter, *J. Mol. Biol.*, 227: 381 (1991); Marks *et al.*, *J. Mol. Biol.*, 222: 581 (1991). The techniques of Cole *et al.* and Boerner *et al.* are also available for the preparation of human monoclonal antibodies. Cole *et al.*, *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, p. 77 (1985) and Boerner *et al.*, *J. Immunol.*, 147(1): 86-95 (1991). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, *e.g.*, mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed that closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; and 5,661,016, and in the following scientific publications: Marks *et al.*, *Bio/Technology*, 10: 779-783 (1992); Lonberg *et al.*, *Nature*, 368: 856-859 (1994); Morrison, *Nature*, 368: 812-813 (1994); Fishwild *et al.*, *Nature Biotechnology*, 14: 845-851 (1996); Neuberger, *Nature Biotechnology*, 14: 826 (1996); Lonberg and Huszar, *Intern. Rev. Immunol.*, 13: 65-93 (1995).

#### 5.2.5.4. Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for the PRO polypeptide, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities. Milstein and Cuello, *Nature*, 305: 537-539 (1983). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, *EMBO J.*, 10: 3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant-domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further

details of generating bispecific antibodies, *see*, for example, Suresh *et al.*, Methods in Enzymology, 121: 210 (1986).

#### 5.2.5.5. Heteroconjugate Antibodies

Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune-system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection. WO 91/00360; WO 92/200373; EP 03089. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide-exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

#### 5.2.5.6. Effector Function Engineering

It may be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, *e.g.*, the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) may be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). *See*, Caron *et al.*, J. Exp. Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff *et al.*, Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. *See*, Stevenson *et al.*, Anti-Cancer Drug Design, 3: 219-230 (1989).

#### 5.2.5.7. Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (*e.g.*, an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (*i.e.*, a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcun, croton, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include <sup>212</sup>Bi, <sup>131</sup>I, <sup>131</sup>In, <sup>90</sup>Y, and <sup>186</sup>Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives



of imidoesters (such as dimethyl adipimidate HCl), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta *et al.*, Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See, WO94/11026.

In another embodiment, the antibody may be conjugated to a "receptor" (such as streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is conjugated to a cytotoxic agent (e.g., a radionucleotide).

#### 5.2.5.8. Immunoliposomes

The antibodies disclosed herein may also be formulated as immunoliposomes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein *et al.*, Proc. Natl. Acad. Sci. USA, 82: 3688 (1985); Hwang *et al.*, Proc. Natl. Acad. Sci. USA, 77: 4030 (1980); and U.S. Pat. Nos. 4,485,045 and 4,544,545. Liposomes with enhanced circulation time are disclosed in U.S. Patent No. 5,013,556.

Particularly useful liposomes can be generated by the reverse-phase evaporation method with a lipid composition comprising phosphatidylcholine, cholesterol, and PEG-derivatized phosphatidylethanolamine (PEG-PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter. Fab' fragments of the antibody of the present invention can be conjugated to the liposomes as described in Martin *et al.*, J. Biol. Chem., 257: 286-288 (1982) via a disulfide-interchange reaction. A chemotherapeutic agent (such as Doxorubicin) is optionally contained within the liposome. See, Gabizon *et al.*, J. National Cancer Inst., 81(19): 1484 (1989).

#### 5.2.5.9. Pharmaceutical Compositions of Antibodies

Antibodies specifically binding a PRO polypeptide identified herein, as well as other molecules identified by the screening assays disclosed hereinbefore, can be administered for the treatment of various disorders as noted above and below in the form of pharmaceutical compositions.

If the PRO polypeptide is intracellular and whole antibodies are used as inhibitors, internalizing antibodies are preferred. However, lipofections or liposomes can also be used to deliver the antibody, or an antibody fragment, into cells. Where antibody fragments are used, the smallest inhibitory fragment that specifically binds to the binding domain of the target protein is preferred. For example, based upon the variable-region sequences of an antibody, peptide molecules can be designed that retain the ability to bind the target protein sequence. Such peptides can be synthesized chemically and/or produced by recombinant DNA technology. See, e.g., Marasco *et al.*, Proc. Natl. Acad. Sci. USA, 90: 7889-7893 (1993).

The formulation herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. Alternatively, or in addition, the composition may comprise an agent that enhances its function, such as, for example, a cytotoxic agent, cytokine, chemotherapeutic agent, or growth-inhibitory agent. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

The active ingredients may also be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methacrylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles, and nanocapsules) or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences, *supra*.

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

Sustained-release preparations may be prepared. Suitable examples of sustained-release preparations include semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, *e.g.*, films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and  $\gamma$  ethyl-L-glutamate, non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOT<sup>TM</sup> (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(-)-3-hydroxybutyric acid. While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods. When encapsulated antibodies remain in the body for a long time, they may denature or aggregate as a result of exposure to moisture at 37°C, resulting in a loss of biological activity and possible changes in immunogenicity. Rational strategies can be devised for stabilization depending on the mechanism involved. For example, if the aggregation mechanism is discovered to be intermolecular S-S bond formation through thio-disulfide interchange, stabilization may be achieved by modifying sulfhydryl residues, lyophilizing from acidic solutions, controlling moisture content, using appropriate additives, and developing specific polymer matrix compositions.

#### 5.2.5.10. Methods of Treatment using the Antibody

It is contemplated that the antibodies to a PRO polypeptide may be used to treat various cardiovascular, endothelial, and angiogenic conditions as noted above.

The antibodies are administered to a mammal, preferably a human, in accord with known methods, such as intravenous administration as a bolus or by continuous infusion over a period of time, by intramuscular, intraperitoneal, intracerebrospinal, subcutaneous, intra-articular, intrasynovial, intrathecal, oral, topical, or inhalation routes. Intravenous administration of the antibody is preferred.

Other therapeutic regimens may be combined with the administration of the antibodies of the instant invention as noted above. For example, if the antibodies are to treat cancer, the patient to be treated with such

antibodies may also receive radiation therapy. Alternatively, or in addition, a chemotherapeutic agent may be administered to the patient. Preparation and dosing schedules for such chemotherapeutic agents may be used according to manufacturers' instructions or as determined empirically by the skilled practitioner. Preparation and dosing schedules for such chemotherapy are also described in Chemotherapy Service, Ed., M.C. Perry (Williams & Wilkins: Baltimore, MD, 1992). The chemotherapeutic agent may precede, or follow administration of the antibody, or may be given simultaneously therewith. The antibody may be combined with an anti-estrogen compound such as tamoxifen or EVISTA™ or an anti-progesterone such as onapristone (*see*, EP 616812) in dosages known for such molecules.

If the antibodies are used for treating cancer, it may be desirable also to administer antibodies against other tumor-associated antigens, such as antibodies that bind to one or more of the ErbB2, EGFR, ErbB3, ErbB4, or VEGF receptor(s). These also include the agents set forth above. Also, the antibody is suitably administered serially or in combination with radiological treatments, whether involving irradiation or administration of radioactive substances. Alternatively, or in addition, two or more antibodies binding the same or two or more different antigens disclosed herein may be co-administered to the patient. Sometimes, it may be beneficial also to administer one or more cytokines to the patient. In a preferred embodiment, the antibodies herein are co-administered with a growth-inhibitory agent. For example, the growth-inhibitory agent may be administered first, followed by an antibody of the present invention. However, simultaneous administration or administration of the antibody of the present invention first is also contemplated. Suitable dosages for the growth-inhibitory agent are those presently used and may be lowered due to the combined action (synergy) of the growth-inhibitory agent and the antibody herein.

In one embodiment, vascularization of tumors is attacked in combination therapy. The anti-PRO polypeptide antibody and another antibody (*e.g.*, anti-VEGF) are administered to tumor-bearing patients at therapeutically effective doses as determined, for example, by observing necrosis of the tumor or its metastatic foci, if any. This therapy is continued until such time as no further beneficial effect is observed or clinical examination shows no trace of the tumor or any metastatic foci. Then TNF is administered, alone or in combination with an auxiliary agent such as alpha-, beta-, or gamma-interferon, anti-HER2 antibody, heregulin, anti-heregulin antibody, D-factor, interleukin-1 (IL-1), interleukin-2 (IL-2), granulocyte-macrophage colony stimulating factor (GM-CSF), or agents that promote microvascular coagulation in tumors, such as anti-protein C antibody, anti-protein S antibody, or C4b binding protein (*see*, WO 91/01753, published 21 February 1991), or heat or radiation.

Since the auxiliary agents will vary in their effectiveness, it is desirable to compare their impact on the tumor by matrix screening in conventional fashion. The administration of anti-PRO polypeptide antibody and TNF is repeated until the desired clinical effect is achieved. Alternatively, the anti-PRO polypeptide antibody is administered together with TNF and, optionally, auxiliary agent(s). In instances where solid tumors are found in the limbs or in other locations susceptible to isolation from the general circulation, the therapeutic agents described herein are administered to the isolated tumor or organ. In other embodiments, a FGF or PDGF antagonist, such as an anti-FGF or an anti-PDGF neutralizing antibody, is administered to the patient in conjunction with the anti-PRO

polypeptide antibody. Treatment with anti-PRO polypeptide antibodies preferably may be suspended during periods of wound healing or desirable neovascularization.

For the prevention or treatment of cardiovascular, endothelial, and angiogenic disorder, the appropriate dosage of an antibody herein will depend on the type of disorder to be treated, as defined above, the severity and course of the disease, whether the antibody is administered for preventive or therapeutic purposes, previous therapy, the patient's clinical history and response to the antibody, and the discretion of the attending physician. The antibody is suitably administered to the patient at one time or over a series of treatments.

For example, depending on the type and severity of the disorder, about 1  $\mu$ g/kg to 50 mg/kg (e.g., 0.1-20 mg/kg) of antibody is an initial candidate dosage for administration to the patient, whether, for example, by one or more separate administrations, or by continuous infusion. A typical daily or weekly dosage might range from about 1  $\mu$ g/kg to 100 mg/kg or more, depending on the factors mentioned above. For repeated administrations over several days or longer, depending on the condition, the treatment is repeated or sustained until a desired suppression of disorder symptoms occurs. However, other dosage regimens may be useful. The progress of this therapy is easily monitored by conventional techniques and assays, including, for example, radiographic tumor imaging.

#### 5.2.5.11. Articles of Manufacture with Antibodies

An article of manufacture containing a container with the antibody and a label is also provided. Such articles are described above, wherein the active agent is an anti-PRO antibody.

#### 5.2.5.12. Diagnosis and Prognosis of Tumors using Antibodies

If the indication for which the antibodies are used is cancer, while cell-surface proteins, such as growth receptors over expressed in certain tumors, are excellent targets for drug candidates or tumor (e.g., cancer) treatment, the same proteins along with PRO polypeptides find additional use in the diagnosis and prognosis of tumors. For example, antibodies directed against the PRO polypeptides may be used as tumor diagnostics or prognostics.

For example, antibodies, including antibody fragments, can be used qualitatively or quantitatively to detect the expression of genes including the gene encoding the PRO polypeptide. The antibody preferably is equipped with a detectable, e.g., fluorescent label, and binding can be monitored by light microscopy, flow cytometry, fluorimetry, or other techniques known in the art. Such binding assays are performed essentially as described above.

*In situ* detection of antibody binding to the marker gene products can be performed, for example, by immunofluorescence or immunoelectron microscopy. For this purpose, a histological specimen is removed from the patient, and a labeled antibody is applied to it, preferably by overlaying the antibody on a biological sample. This procedure also allows for determining the distribution of the marker gene product in the tissue examined. It will be apparent to those skilled in the art that a wide variety of histological methods are readily available for *in situ* detection.

The following Examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

The disclosures of all patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

5        6.        EXAMPLES

Commercially available reagents referred to in the Examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following Examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Manassas, VA. Unless otherwise noted, the present invention uses standard procedures of recombinant DNA technology, such as those described hereinabove and in the following textbooks: Sambrook *et al.*, *supra*; Ausubel *et al.*, Current Protocols in Molecular Biology (Green Publishing Associates and Wiley Interscience, N.Y., 1989); Innis *et al.*, PCR Protocols: A Guide to Methods and Applications (Academic Press, Inc.: N.Y., 1990); Harlow *et al.*, Antibodies: A Laboratory Manual (Cold Spring Harbor Press: Cold Spring Harbor, 1988); Gait, Oligonucleotide Synthesis (IRL Press: Oxford, 1984); Freshney, Animal Cell Culture, 1987; Coligan *et al.*, Current Protocols in Immunology, 1991.

15            6.1.        EXAMPLE 1: Extracellular Domain Homology Screening to Identify Novel Polypeptides and cDNA Encoding Therefor

The extracellular domain (ECD) sequences (including the secretion signal sequence, if any) from about 950 known secreted proteins from the Swiss-Prot public database were used to search EST databases. The EST databases included public databases (*e.g.*, Dayhoff, GenBank), and proprietary databases (*e.g.* LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST-2 (Altschul *et al.*, Methods in Enzymology, 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequences. Those comparisons with a BLAST score of 70 (or in some cases, 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, WA).

25            Using this extracellular domain homology screen, consensus DNA sequences were assembled relative to the other identified EST sequences using phrap. In addition, the consensus DNA sequences obtained were often (but not always) extended using repeated cycles of BLAST or BLAST-2 and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above.

30            Based upon the consensus sequences obtained as described above, oligonucleotides were then synthesized and used to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for a PRO polypeptide. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5 kbp. In order to screen several libraries for a full-length clone, 35        DNA from the libraries was screened by PCR amplification, as per Ausubel *et al.*, Current Protocols in Molecular

Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; *see*, Holmes *et al.*, *Science*, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

## 6.2. EXAMPLE 2: Isolation of cDNA Clones by Amylase Screening

### 6.2.1. Preparation of oligo dT primed cDNA library

mRNA was isolated from a human tissue of interest using reagents and protocols from Invitrogen, San Diego, CA (Fast Track 2). This RNA was used to generate an oligo dT primed cDNA library in the vector pRK5D using reagents and protocols from Life Technologies, Gaithersburg, MD (Super Script Plasmid System). In this procedure, the double stranded cDNA was sized to greater than 1000 bp and the SalI/NotI linked cDNA was cloned into XhoI/NotI cleaved vector. pRK5D is a cloning vector that has an sp6 transcription initiation site followed by an SfiI restriction enzyme site preceding the XhoI/NotI cDNA cloning sites.

### 6.2.2. Preparation of random primed cDNA library

A secondary cDNA library was generated in order to preferentially represent the 5' ends of the primary cDNA clones. Sp6 RNA was generated from the primary library (described above), and this RNA was used to generate a random primed cDNA library in the vector pSST-AMY.0 using reagents and protocols from Life Technologies (Super Script Plasmid System, referenced above). In this procedure the double stranded cDNA was sized to 500-1000 bp, linked with blunt to NotI adaptors, cleaved with SfiI, and cloned into SfiI/NotI cleaved vector. pSST-AMY.0 is a cloning vector that has a yeast alcohol dehydrogenase promoter preceding the cDNA cloning sites and the mouse amylase sequence (the mature sequence without the secretion signal) followed by the yeast alcohol dehydrogenase terminator, after the cloning sites. Thus, cDNAs cloned into this vector that are fused in frame with amylase sequence will lead to the secretion of amylase from appropriately transfected yeast colonies.

### 6.2.3. Transformation and Detection

DNA from the library described in paragraph 2 above was chilled on ice to which was added electrocompetent DH10B bacteria (Life Technologies, 20 ml). The bacteria and vector mixture was then electroporated as recommended by the manufacturer. Subsequently, SOC media (Life Technologies, 1 ml) was added and the mixture was incubated at 37°C for 30 minutes. The transformants were then plated onto 20 standard 150 mm LB plates containing ampicillin and incubated for 16 hours (37°C). Positive colonies were scraped off the plates and the DNA was isolated from the bacterial pellet using standard protocols, *e.g.*, CsCl-gradient. The purified DNA was then carried on to the yeast protocols below.

The yeast methods were divided into three categories: (1) Transformation of yeast with the plasmid/cDNA combined vector; (2) Detection and isolation of yeast clones secreting amylase; and (3) PCR amplification of the insert directly from the yeast colony and purification of the DNA for sequencing and further analysis.

5 The yeast strain used was HD56-5A (ATCC-90785). This strain has the following genotype: MAT alpha, ura3-52, leu2-3, leu2-112, his3-11, his3-15, MAL<sup>+</sup>, SUC<sup>+</sup>, GAL<sup>+</sup>. Preferably, yeast mutants can be employed that have deficient post-translational pathways. Such mutants may have translocation deficient alleles in *sec71*, *sec72*, *sec62*, with truncated *sec71* being most preferred. Alternatively, antagonists (including antisense nucleotides and/or ligands) which interfere with the normal operation of these genes, other proteins implicated in this post translation pathway (e.g., SEC61p, SEC72p, SEC62p, SEC63p, TDJ1p or SSA1p-4p) or the complex formation  
10 of these proteins may also be preferably employed in combination with the amylase-expressing yeast.

Transformation was performed based on the protocol outlined by Gietz *et al.*, Nucl. Acid. Res., 20:1425 (1992). Transformed cells were then inoculated from agar into YEPD complex media broth (100 ml) and grown overnight at 30°C. The YEPD broth was prepared as described in Kaiser *et al.*, Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 207 (1994). The overnight culture was then diluted to about 2  
15 x 10<sup>6</sup> cells/ml (approx. OD<sub>600</sub>=0.1) into fresh YEPD broth (500 ml) and regrown to 1 x 10<sup>7</sup> cells/ml (approx. OD<sub>600</sub>=0.4-0.5).

The cells were then harvested and prepared for transformation by transfer into GS3 rotor bottles in a Sorval GS3 rotor at 5,000 rpm for 5 minutes, the supernatant discarded, and then resuspended into sterile water, and centrifuged again in 50 ml falcon tubes at 3,500 rpm in a Beckman GS-6KR centrifuge. The supernatant was  
20 discarded and the cells were subsequently washed with LiAc/TE (10 ml, 10 mM Tris-HCl, 1 mM EDTA pH 7.5, 100 mM Li<sub>2</sub>OOCCH<sub>3</sub>), and resuspended into LiAc/TE (2.5 ml).

Transformation took place by mixing the prepared cells (100 µl) with freshly denatured single stranded salmon testes DNA (Lofstrand Labs, Gaithersburg, MD) and transforming DNA (1 µg, vol. < 10 µl) in microfuge tubes. The mixture was mixed briefly by vortexing, then 40% PEG/TE (600 µl, 40% polyethylene glycol-4000,  
25 10 mM Tris-HCl, 1 mM EDTA, 100 mM Li<sub>2</sub>OOCCH<sub>3</sub>, pH 7.5) was added. This mixture was gently mixed and incubated at 30°C while agitating for 30 minutes. The cells were then heat shocked at 42°C for 15 minutes, and the reaction vessel centrifuged in a microfuge at 12,000 rpm for 5-10 seconds, decanted and resuspended into TE (500 µl, 10 mM Tris-HCl, 1 mM EDTA pH 7.5) followed by recentrifugation. The cells were then diluted into TE (1 ml) and aliquots (200 µl) were spread onto the selective media previously prepared in 150 mm growth plates  
30 (VWR).

Alternatively, instead of multiple small reactions, the transformation was performed using a single, large scale reaction, wherein reagent amounts were scaled up accordingly.

The selective media used was a synthetic complete dextrose agar lacking uracil (SCD-Ura) prepared as described in Kaiser *et al.*, Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 208-  
35 210 (1994). Transformants were grown at 30°C for 2-3 days.

The detection of colonies secreting amylase was performed by including red starch in the selective growth media. Starch was coupled to the red dye (Reactive Red-120, Sigma) as per the procedure described by Biely *et*

*al.*, Anal. Biochem., 172:176-179 (1988). The coupled starch was incorporated into the SCD-Ura agar plates at a final concentration of 0.15% (w/v), and was buffered with potassium phosphate to a pH of 7.0 (50-100 mM final concentration).

5 The positive colonies were picked and streaked across fresh selective media (onto 150 mm plates) in order to obtain well isolated and identifiable single colonies. Well isolated single colonies positive for amylase secretion were detected by direct incorporation of red starch into buffered SCD-Ura agar. Positive colonies were determined by their ability to break down starch resulting in a clear halo around the positive colony visualized directly.

#### 6.2.4. Isolation of DNA by PCR Amplification

10 When a positive colony was isolated, a portion of it was picked by a toothpick and diluted into sterile water (30  $\mu$ l) in a 96 well plate. At this time, the positive colonies were either frozen and stored for subsequent analysis or immediately amplified. An aliquot of cells (5  $\mu$ l) was used as a template for the PCR reaction in a 25  $\mu$ l volume containing: 0.5  $\mu$ l KlenTaq (Clontech, Palo Alto, CA); 4.0  $\mu$ l 10 mM dNTP's (Perkin Elmer-Cetus); 2.5  $\mu$ l KlenTaq buffer (Clontech); 0.25  $\mu$ l forward oligo 1; 0.25  $\mu$ l reverse oligo 2; 12.5  $\mu$ l distilled water. The sequence of the forward oligonucleotide 1 was:

15 5'-TGTAACGACGCGCCAGTTAAATAGACCTGCAATTATTAATCT-3' (SEQ ID NO:382)

The sequence of reverse oligonucleotide 2 was:

5'-CAGGAAACAGCTATGACCACCTGCACACCTGCAAATCCATT-3' (SEQ ID NO:383)

PCR was then performed as follows:

	a.	Denature	92°C, 5 minutes
20	b.	3 cycles of:	Denature
			92°C, 30 seconds
			Anneal
			59°C, 30 seconds
			Extend
			72°C, 60 seconds
	c.	3 cycles of:	Denature
			92°C, 30 seconds
25			Anneal
			57°C, 30 seconds
			Extend
			72°C, 60 seconds
	d.	25 cycles of:	Denature
			92°C, 30 seconds
			Anneal
			55°C, 30 seconds
			Extend
			72°C, 60 seconds
	e.	Hold	4°C

30 The underlined regions of the oligonucleotides annealed to the ADH promoter region and the amylase region, respectively, and amplified a 307 bp region from vector pSST-AMY.0 when no insert was present. Typically, the first 18 nucleotides of the 5' end of these oligonucleotides contained annealing sites for the sequencing primers. Thus, the total product of the PCR reaction from an empty vector was 343 bp. However, signal sequence-fused cDNA resulted in considerably longer nucleotide sequences.



Following the PCR, an aliquot of the reaction (5 µl) was examined by agarose gel electrophoresis in a 1% agarose gel using a Tris-Borate-EDTA (TBE) buffering system as described by Sambrook *et al.*, *supra*. Clones resulting in a single strong PCR product larger than 400 bp were further analyzed by DNA sequencing after purification with a 96 Qiaquick PCR clean-up column (Qiagen Inc., Chatsworth, CA).

5           6.3.    EXAMPLE 3: Isolation of cDNA Clones Using Signal Algorithm Analysis

Various polypeptide-encoding nucleic acid sequences were identified by applying a proprietary signal sequence finding algorithm developed by Genentech, Inc., (South San Francisco, CA) upon ESTs as well as clustered and assembled EST fragments from public (*e.g.*, GenBank) and/or private (LIFESEQ®, Incyte Pharmaceuticals, Inc., Palo Alto, CA) databases. The signal sequence algorithm computes a secretion signal score based on the character of the DNA nucleotides surrounding the first and optionally the second methionine codon(s) (ATG) at the 5'-end of the sequence or sequence fragment under consideration. The nucleotides following the first ATG must code for at least 35 unambiguous amino acids without any stop codons. If the first ATG has the required amino acids, the second is not examined. If neither meets the requirement, the candidate sequence is not scored. In order to determine whether the EST sequence contains an authentic signal sequence, the DNA and corresponding amino acid sequences surrounding the ATG codon are scored using a set of seven sensors (evaluation parameters) known to be associated with secretion signals. Use of this algorithm resulted in the identification of numerous polypeptide-encoding nucleic acid sequences.

6.4.    EXAMPLE 4: Isolation of cDNA clones Encoding Human PRO Polypeptides

Using the techniques described in Examples 1 to 3 above, numerous full-length cDNA clones were identified as encoding PRO polypeptides as disclosed herein. These cDNAs were then deposited under the terms of the Budapest Treaty with the American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209, USA (ATCC) as shown in Table 7 below.

Table 7

<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
23330-1390	209775	4/14/1998
23339-1130	209282	9/18/1997
26846-1397	203406	10/27/1998
26847-1395	209772	4/14/1998
27865-1091	209296	9/23/1997
30868-1156	1437-PTA	3/2/2000
30871-1157	209380	10/16/1997
32286-1191	209385	10/16/1997
33089-1132	209262	9/16/1997
33092-1202	209420	10/28/1997

	33100-1159	209377	10/16/1997
	33223-1136	209264	9/16/1997
	34392-1170	209526	12/10/1997
	34431-1177	209399	10/17/1997
5	34433-1308	209719	3/31/1998
	34434-1139	209252	9/16/1997
	35600-1162	209370	10/16/1997
	35673-1201	209418	10/28/1997
	35880-1160	209379	10/16/1997
10	35918-1174	209402	10/17/1997
	36350-1158	209378	10/16/1997
	36638-1056	209456	11/12/1997
	38268-1188	209421	10/28/1997
	40370-1217	209485	11/21/1997
15	40628-1216	209432	11/7/1997
	43316-1237	209487	11/21/1997
	44196-1353	209847	5/6/1998
	45409-2511	203579	1/12/1999
	45419-1252	209616	2/5/1998
20	46777-1253	209619	2/5/1998
	48336-1309	209669	3/11/1998
	48606-1479	203040	7/1/1998
	49435-1219	209480	11/21/1997
	49631-1328	209806	4/28/1998
25	50919-1361	209848	5/6/1998
	50920-1325	209700	3/26/1998
	50921-1458	209859	5/12/1998
	52758-1399	209773	4/14/1998
	53517-1366-1	209802	4/23/1998
30	53915-1258	209593	1/21/1998
	53974-1401	209774	4/14/1998
	53987-1438	209858	5/12/1998
	56047-1456	209948	6/9/1998
	56050-1455	203011	6/23/1998
35	56110-1437	203113	8/11/1998
	56405-1357	209849	5/6/1998
	56433-1406	209857	5/12/1998

	56439-1376	209864	5/14/1998
	56529-1647	203293	9/29/1998
	56865-1491	203022	6/23/1998
	56965-1356	209842	5/6/1998
5	57033-1403-1	209905	5/27/1998
	57037-1444	209903	5/27/1998
	57039-1402	209777	4/14/1998
	57689-1385	209869	5/14/1998
	57690-1374	209950	6/9/1998
10	57694-1341	203017	6/23/1998
	57695-1340	203006	6/23/1998
	57699-1412	203020	6/23/1998
	57700-1408	203583	1/12/1999
	57708-1411	203021	6/23/1998
15	57838-1337	203014	6/23/1998
	58847-1383	209879	5/20/1998
	58852-1637	203271	9/22/1998
	58853-1423	203016	6/23/1998
	59212-1627	203245	9/9/1998
20	59220-1514	209962	6/9/1998
	59493-1420	203050	7/1/1998
	59497-1496	209941	6/4/1998
	59586-1520	203288	9/29/1998
	59588-1571	203106	8/11/1998
25	59620-1463	209989	6/16/1998
	59622-1334	209984	6/16/1998
	59777-1480	203111	8/11/1998
	59848-1512	203088	8/4/1998
	59849-1504	209986	6/16/1998
30	60621-1516	203091	8/4/1998
	60622-1525	203090	8/4/1998
	60764-1533	203452	11/10/1998
	60783-1611	203130	8/18/1998
	61755-1554	203112	8/11/1998
35	62306-1570	203254	9/9/1998
	62312-2558	203836	3/9/1999
	62814-1521	203093	8/4/1998

	62872-1509	203100	8/4/1998
	64883-1526	203253	9/9/1998
	64886-1601	203241	9/9/1998
	64889-1541	203250	9/9/1998
5	64896-1539	203238	9/9/1998
	64897-1628	203216	9/15/1998
	64903-1553	203223	9/15/1998
	64908-1163-1	203243	9/9/1998
	64950-1590	203224	9/15/1998
10	65402-1540	203252	9/9/1998
	65404-1551	203244	9/9/1998
	65405-1547	203476	11/17/1998
	65410-1569	203231	9/15/1998
	65412-1523	203094	8/4/1998
15	66307-2661	431-PTA	7/27/1999
	66526-1616	203246	9/9/1998
	66659-1593	203269	9/22/1998
	66660-1585	203279	9/22/1998
	66667-1596	203267	9/22/1998
20	66672-1586	203265	9/22/1998
	66675-1587	203282	9/22/1998
	67300-1605	203163	8/25/1998
	68818-2536	203657	2/9/1999
	68862-2546	203652	2/9/1999
25	68872-1620	203160	8/25/1998
	71290-1630	203275	9/22/1998
	73736-1657	203466	11/17/1998
	73739-1645	203270	9/22/1998
	73742-1662	203316	10/6/1998
30	76385-1692	203664	2/9/1999
	76393-1664	203323	10/6/1998
	76399-1700	203472	11/17/1998
	76400-2528	203573	1/12/1999
	76510-2504	203477	11/17/1998
35	76529-1666	203315	10/6/1998
	76532-1702	203473	11/17/1998
	76541-1675	203409	10/27/1998

	77503-1686	203362	10/20/1998
	77624-2515	203553	12/22/1998
	79230-2525	203549	12/22/1998
	79862-2522	203550	12/22/1998
5	80145-2594	204-PTA	6/8/1999
	80899-2501	203539	12/15/1998
	81754-2532	203542	12/15/1998
	81757-2512	203543	12/15/1998
	81761-2583	203862	3/23/1999
10	82358-2738	510-PTA	8/10/1999
	82364-2538	203603	1/20/1999
	82403-2959	2317-PTA	8/1/2000
	83500-2506	203391	10/29/1998
	83560-2569	203816	3/2/1999
15	84210-2576	203818	3/2/1999
	84920-2614	203966	4/27/1999
	86576-2595	203868	3/23/1999
	92218-2554	203834	3/9/1999
	92233-2599	134-PTA	5/25/1999
20	92256-2596	203891	3/30/1999
	92265-2669	256-PTA	6/22/1999
	92274-2617	203971	4/27/1999
	92929-2534-1	203586	1/12/1999
	93011-2637	20-PTA	5/4/1999
25	94854-2586	203864	3/23/1999
	96787-2534-1	203589	1/12/1999
	96867-2620	203972	4/27/1999
	96872-2674	550-PTA	8/17/1999
	96878-2626	23-PTA	5/4/1999
30	96889-2641	119-PTA	5/25/1999
	100312-2645	44-PTA	5/11/1999
	105782-2693	387-PTA	7/20/1999
	105849-2704	473-PTA	8/3/1999
	108725-2766	863-PTA	10/19/1999
35	108769-2765	861-PTA	10/19/1999
	119498-2965	2298-PTA	7/25/2000
	119535-2756	613-PTA	8/31/1999

	125185-2806	1031-PTA	12/7/1999
	131639-2874	1784-PTA	4/25/2000
	139623-2893	1670-PTA	4/11/2000
	143076-2787	1028-PTA	12/7/1999
5	143276-2975	2387-PTA	8/8/2000
	164625-2890	1535-PTA	3/21/2000
	167678-2963	2302-PTA	7/25/2000
	170021-2923	1906-PTA	5/23/2000
	170212-3000	2583-PTA	10/10/2000
10	177313-2982	2251-PTA	7/19/2000

These deposits were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposits will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC § 122 and the Commissioner's rules pursuant thereto (including 37 CFR § 1.14 with particular reference to 886 OG 638).

The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

**6.5 EXAMPLE 5: Isolation of cDNA clones Encoding Human PRO1873, PRO7223, PRO7248, PRO730, PRO532, PRO7261, PRO734, PRO771, PRO2010, PRO5723, PRO3444, PRO9940, PRO3562, PRO10008, PRO5730, PRO6008, PRO4527, PRO4538 and PRO4553**

DNA molecules encoding the PRO1873, PRO7223, PRO7248, PRO730, PRO532, PRO7261, PRO734, PRO771, PRO2010, PRO5723, PRO3444, PRO9940, PRO3562, PRO10008, PRO5730, PRO6008, PRO4527, PRO4538 and PRO4553 polypeptides shown in the accompanying figures were obtained through GenBank.

**6.6. EXAMPLE 6: Use of PRO as a Hybridization Probe**

The following method describes use of a nucleotide sequence encoding PRO as a hybridization probe.

DNA comprising the coding sequence of full-length or mature PRO (as shown in accompanying figures) or a fragment thereof is employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of PRO) in human tissue cDNA libraries or human tissue genomic libraries.

5 Hybridization and washing of filters containing either library DNAs is performed under the following high-stringency conditions. Hybridization of radiolabeled probe derived from the gene encoding PRO polypeptide to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

10 DNAs having a desired sequence identity with the DNA encoding full-length native sequence can then be identified using standard techniques known in the art.

#### 6.7. EXAMPLE 7: Expression of PRO in *E. coli*

This example illustrates preparation of an unglycosylated form of PRO by recombinant expression in *E. coli*.

15 The DNA sequence encoding PRO is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see, Bolivar *et al.*, Gene, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a poly-His leader (including the first six STII codons, poly-His sequence, and enterokinase cleavage site),  
20 the PRO coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook *et al.*, *supra*. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA  
25 sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

30 After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO protein can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

PRO may be expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding PRO is initially amplified using selected PCR primers. The primers will contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences  
35 providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences are then ligated into an

expression vector, which is used to transform an *E. coli* host based on strain 52 (W3110 fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq). Transformants are first grown in LB containing 50 mg/ml carbenicillin at 30°C with shaking until an OD<sub>600</sub> of 3-5 is reached. Cultures are then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.71 g sodium citrate·2H<sub>2</sub>O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g Sheffield hycase SF in 500 ml water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM MgSO<sub>4</sub>) and grown for approximately 20-30 hours at 30°C with shaking. Samples are removed to verify expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets are frozen until purification and refolding.

*E. coli* paste from 0.5 to 1 L fermentations (6-10 g pellets) is resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution is stirred overnight at 4°C. This step results in a denatured protein with all cysteine residues blocked by sulfitolization. The solution is centrifuged at 40,000 rpm in a Beckman Ultracentrifuge for 30 min. The supernatant is diluted with 3-5 volumes of metal chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. The clarified extract is loaded onto a 5 ml Qiagen Ni<sup>2+</sup>-NTA metal chelate column equilibrated in the metal chelate column buffer. The column is washed with additional buffer containing 50 mM imidazole (Calbiochem, Utrol grade), pH 7.4. The protein is eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein are pooled and stored at 4°C. Protein concentration is estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

The proteins are refolded by diluting the sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM cysteine, 20 mM glycine and 1 mM EDTA. Refolding volumes are chosen so that the final protein concentration is between 50 to 100 micrograms/ml. The refolding solution is stirred gently at 4°C for 12-36 hours. The refolding reaction is quenched by the addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further purification of the protein, the solution is filtered through a 0.22 micron filter and acetonitrile is added to 2-10% final concentration. The refolded protein is chromatographed on a Poros R1/H reversed phase column using a mobile buffer of 0.1% TFA with elution with a gradient of acetonitrile from 10 to 80%. Aliquots of fractions with A<sub>280</sub> absorbance are analyzed on SDS polyacrylamide gels and fractions containing homogeneous refolded protein are pooled. Generally, the properly refolded species of most proteins are eluted at the lowest concentrations of acetonitrile since those species are the most compact with their hydrophobic interiors shielded from interaction with the reversed phase resin. Aggregated species are usually eluted at higher acetonitrile concentrations. In addition to resolving misfolded forms of proteins from the desired form, the reversed phase step also removes endotoxin from the samples.

Fractions containing the desired folded PRO polypeptide are pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins are formulated into 20 mM Hepes, pH 6.8 with 0.14 M sodium chloride and 4% mannitol by dialysis or by gel filtration using G25 Superfine (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.



6.8. EXAMPLE 8: Expression of PRO in mammalian cells

This example illustrates preparation of a potentially glycosylated form of PRO by recombinant expression in mammalian cells.

5 The vector, pRK5 (*see* EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the PRO DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO DNA using ligation methods such as described in Sambrook *et al.*, *supra*. The resulting vector is called pRK5-PRO.

10 In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10  $\mu$ g pRK5-PRO DNA is mixed with about 1  $\mu$ g DNA encoding the VA RNA gene [Thimmappaya *et al.*, *Cell*, 31:543 (1982)] and dissolved in 500  $\mu$ l of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M  $\text{CaCl}_2$ . To this mixture is added, dropwise, 500  $\mu$ l of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM  $\text{NaPO}_4$ , and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, 15 fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200  $\mu$ Ci/ml  $^{35}\text{S}$ -cysteine and 200  $\mu$ Ci/ml  $^{35}\text{S}$ -methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of the PRO polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays. 20

In an alternative technique, PRO may be introduced into 293 cells transiently using the dextran sulfate method described by Sompanyrac *et al.*, *Proc. Natl. Acad. Sci.*, 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700  $\mu$ g pRK5-PRO DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5  $\mu$ g/ml bovine insulin and 0.1  $\mu$ g/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography. 25 30

In another embodiment, PRO can be expressed in CHO cells. The pRK5-PRO can be transfected into CHO cells using known reagents such as  $\text{CaPO}_4$  or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as  $^{35}\text{S}$ -methionine. After determining the presence of a PRO polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. 35

The medium containing the expressed PRO polypeptide can then be concentrated and purified by any selected method.

Epitope-tagged PRO may also be expressed in host CHO cells. The PRO may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-His tag into a Baculovirus expression vector. The poly-His tagged PRO insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO can then be concentrated and purified by any selected method, such as by Ni<sup>2+</sup>-chelate affinity chromatography.

PRO may also be expressed in CHO and/or COS cells by a transient expression procedure or in CHO cells by another stable expression procedure.

Stable expression in CHO cells is performed using the following procedure. The proteins are expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (*e.g.*, extracellular domains) of the respective proteins are fused to an IgG1 constant region sequence containing the hinge, CH2 and CH2 domains and/or as a poly-His tagged form.

Following PCR amplification, the respective DNAs are subcloned in a CHO expression vector using standard techniques as described in Ausubel *et al.*, Current Protocols of Molecular Biology, Unit 3.16, John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used in expression in CHO cells is as described in Lucas *et al.*, Nucl. Acids Res., 24:9 (1774-1779 (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits selection for stable maintenance of the plasmid following transfection.

Twelve micrograms of the desired plasmid DNA is introduced into approximately 10 million CHO cells using commercially available transfection reagents Superfect® (Qiagen), Dosper® or Fugene® (Boehringer Mannheim). The cells are grown as described in Lucas *et al.*, *supra*. Approximately 3 x 10<sup>7</sup> cells are frozen in an ampule for further growth and production as described below.

The ampules containing the plasmid DNA are thawed by placement into a water bath and mixed by vortexing. The contents are pipetted into a centrifuge tube containing 10 ml of media and centrifuged at 1000 rpm for 5 minutes. The supernatant is aspirated and the cells are resuspended in 10 ml of selective media (0.2 µm filtered PS20 with 5% 0.2 µm diafiltered fetal bovine serum). The cells are then aliquoted into a 100 ml spinner containing 90 ml of selective media. After 1-2 days, the cells are transferred into a 250 ml spinner filled with 150 ml selective growth medium and incubated at 37°C. After another 2-3 days, 250 ml, 500 ml and 2000 ml spinners are seeded with 3 x 10<sup>5</sup> cells/ml. The cell media is exchanged with fresh media by centrifugation and resuspension in production medium. Although any suitable CHO media may be employed, a production medium described in U.S. Patent No. 5,122,469, issued June 16, 1992 may actually be used. A 3L production spinner is seeded at 1.2 x 10<sup>6</sup> cells/ml. On day 0, the cell number and pH is determined. On day 1, the spinner is sampled and sparging with filtered air is commenced. On day 2, the spinner is sampled, the temperature shifted to 33°C, and 30 ml of 500 g/L

glucose and 0.6 ml of 10% antifoam (e.g., 35% polydimethylsiloxane emulsion, Dow Corning 365 Medical Grade Emulsion) taken. Throughout the production, the pH is adjusted as necessary to keep it at around 7.2. After 10 days, or until the viability drops below 70%, the cell culture is harvested by centrifugation and filtering through a 0.22  $\mu$ m filter. The filtrate is either stored at 4°C or immediately loaded onto columns for purification.

5 For the poly-His tagged constructs, the proteins are purified using a Ni<sup>2+</sup>-NTA column (Qiagen). Before purification, imidazole is added to the conditioned media to a concentration of 5 mM. The conditioned media is pumped onto a 6 ml Ni<sup>2+</sup>-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column is washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly  
10 purified protein is subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc-containing) constructs are purified from the conditioned media as follows. The conditioned medium is pumped onto a 5 ml Protein A column (Pharmacia) which has been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column is washed extensively with equilibration buffer before  
15 elution with 100 mM citric acid, pH 3.5. The eluted protein is immediately neutralized by collecting 1 ml fractions into tubes containing 275  $\mu$ l of 1 M Tris buffer, pH 9. The highly purified protein is subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity is assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

#### 20 6.9. EXAMPLE 9: Expression of PRO in Yeast

The following method describes recombinant expression of PRO in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO from the ADH2/GAPDH promoter. DNA encoding PRO and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of PRO. For secretion, DNA encoding PRO can be cloned  
25 into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, a native PRO signal peptide or other mammalian signal peptide, or, for example, a yeast alpha-factor or invertase secretory signal/leader sequence, and linker sequences (if needed) for expression of PRO.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by  
30 precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant PRO can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing PRO may further be purified using selected column chromatography resins.

35 Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

6.10. EXAMPLE 10: Expression of PRO in Baculovirus-Infected Insect Cells

The following method describes recombinant expression in Baculovirus-infected insect cells.

The sequence coding for PRO is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-His tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the sequence encoding PRO or the desired portion of the coding sequence of PRO (such as the sequence encoding the extracellular domain of a transmembrane protein or the sequence encoding the mature protein if the protein is extracellular) is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 - 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression are performed as described by O'Reilley *et al.*, Baculovirus expression vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

Expressed poly-His tagged PRO can then be purified, for example, by Ni<sup>2+</sup>-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert *et al.*, Nature, **362**:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 ml Hepes, pH 7.9; 12.5 mM MgCl<sub>2</sub>; 0.1 mM EDTA; 10% glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% glycerol, pH 7.8) and filtered through a 0.45 µm filter. A Ni<sup>2+</sup>-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 ml, washed with 25 ml of water and equilibrated with 25 ml of loading buffer. The filtered cell extract is loaded onto the column at 0.5 ml per minute. The column is washed to baseline A<sub>280</sub> with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A<sub>280</sub> baseline again, the column is developed with a 0 to 500 mM imidazole gradient in the secondary wash buffer. One ml fractions are collected and analyzed by SDS-PAGE and silver staining or Western blot with Ni<sup>2+</sup>-NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His<sub>10</sub>-tagged PRO are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

Following PCR amplification, the respective coding sequences are subcloned into a baculovirus expression vector (pb.PH.IgG for IgG fusions and pb.PH.His.c for poly-His tagged proteins), and the vector and Baculogold® baculovirus DNA (Pharmingen) are co-transfected into 105 *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711), using Lipofectin (Gibco BRL). pb.PH.IgG and pb.PH.His are modifications of the commercially available baculovirus expression vector pVL1393 (Pharmingen), with modified polylinker regions to include the His or Fc tag sequences. The cells are grown in Hink's TNM-FH medium supplemented with 10% FBS (Hyclone). Cells are

incubated for 5 days at 28°C. The supernatant is harvested and subsequently used for the first viral amplification by infecting Sf9 cells in Hink's TNM-FH medium supplemented with 10% FBS at an approximate multiplicity of infection (MOI) of 10. Cells are incubated for 3 days at 28°C. The supernatant is harvested and the expression of the constructs in the baculovirus expression vector is determined by batch binding of 1 ml of supernatant to 25 ml of Ni<sup>2+</sup>-NTA beads (QIAGEN) for histidine tagged proteins or Protein-A Sepharose CL-4B beads (Pharmacia) for IgG tagged proteins followed by SDS-PAGE analysis comparing to a known concentration of protein standard by Coomassie blue staining.

The first viral amplification supernatant is used to infect a spinner culture (500 ml) of Sf9 cells grown in ESF-921 medium (Expression Systems LLC) at an approximate MOI of 0.1. Cells are incubated for 3 days at 28°C. The supernatant is harvested and filtered. Batch binding and SDS-PAGE analysis is repeated, as necessary, until expression of the spinner culture is confirmed.

The conditioned medium from the transfected cells (0.5 to 3 L) is harvested by centrifugation to remove the cells and filtered through 0.22 micron filters. For the poly-His tagged constructs, the protein construct is purified using a Ni<sup>2+</sup>-NTA column (Qiagen). Before purification, imidazole is added to the conditioned media to a concentration of 5 mM. The conditioned media is pumped onto a 6 ml Ni<sup>2+</sup>-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column is washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein is subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc containing) constructs of proteins are purified from the conditioned media as follows. The conditioned media is pumped onto a 5 ml Protein A column (Pharmacia) which has been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column is washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein is immediately neutralized by collecting 1 ml fractions into tubes containing 275 ml of 1 M Tris buffer, pH 9. The highly purified protein is subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity of the proteins is verified by SDS polyacrylamide gel (PAGE) electrophoresis and N-terminal amino acid sequencing by Edman degradation.

Alternatively, a modified baculovirus procedure may be used incorporating high-5 cells. In this procedure, the DNA encoding the desired sequence is amplified with suitable systems, such as Pfu (Stratagene), or fused upstream (5'-of) of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-His tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pIE1-1 (Novagen). The pIE1-1 and pIE1-2 vectors are designed for constitutive expression of recombinant proteins from the baculovirus ie1 promoter in stably-transformed insect cells (1). The plasmids differ only in the orientation of the multiple cloning sites and contain all promoter sequences known to be important for ie1-mediated gene expression in uninfected insect cells as well as the hr5 enhancer element. pIE1-1 and pIE1-2 include the translation initiation site and can be used to produce fusion proteins. Briefly, the desired sequence or the desired portion of the sequence (such as the sequence encoding the

extracellular domain of a transmembrane protein) is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector. For example, derivatives of pIE1-1 can include the Fc region of human IgG (pb.PH.IgG) or an 8 histidine (pb.PH.His) tag downstream (3'-of) the desired sequence. Preferably, the vector construct is sequenced for confirmation.

High-5 cells are grown to a confluency of 50% under the conditions of, 27°C, no CO<sub>2</sub>, NO pen/strep. For each 150 mm plate, 30 µg of pIE based vector containing the sequence is mixed with 1 ml Ex-Cell medium (Media: Ex-Cell 401 + 1/100 L-Glu JRH Biosciences #14401-78P (note: this media is light sensitive)), and in a separate tube, 100 µl of CellFectin (CellFECTIN (GibcoBRL #10362-010) (vortexed to mix)) is mixed with 1 ml of Ex-Cell medium. The two solutions are combined and allowed to incubate at room temperature for 15 minutes. 8 ml of Ex-Cell media is added to the 2 ml of DNA/CellFECTIN mix and this is layered on high-5 cells that have been washed once with Ex-Cell media. The plate is then incubated in darkness for 1 hour at room temperature. The DNA/CellFECTIN mix is then aspirated, and the cells are washed once with Ex-Cell to remove excess CellFECTIN, 30 ml of fresh Ex-Cell media is added and the cells are incubated for 3 days at 28°C. The supernatant is harvested and the expression of the sequence in the baculovirus expression vector is determined by batch binding of 1 ml of supernatant to 25 ml of Ni<sup>2+</sup>-NTA beads (QIAGEN) for histidine tagged proteins or Protein-A Sepharose CL-4B beads (Pharmacia) for IgG tagged proteins followed by SDS-PAGE analysis comparing to a known concentration of protein standard by Coomassie blue staining.

The conditioned media from the transfected cells (0.5 to 3 L) is harvested by centrifugation to remove the cells and filtered through 0.22 micron filters. For the poly-His tagged constructs, the protein comprising the sequence is purified using a Ni<sup>2+</sup>-NTA column (Qiagen). Before purification, imidazole is added to the conditioned media to a concentration of 5 mM. The conditioned media is pumped onto a 6 ml Ni<sup>2+</sup>-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 48°C. After loading, the column is washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein is then subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc containing) constructs of proteins are purified from the conditioned media as follows. The conditioned media is pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column is washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein is immediately neutralized by collecting 1 ml fractions into tubes containing 275 ml of 1 M Tris buffer, pH 9. The highly purified protein is subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity of the sequence is assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation and other analytical procedures as desired or necessary.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

6.11. EXAMPLE 11: Preparation of Antibodies that Bind PRO

This example illustrates preparation of monoclonal antibodies which can specifically bind the PRO polypeptide or an epitope on the PRO polypeptide without substantially binding to any other polypeptide or polypeptide epitope.

5 Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, *supra*. Immunogens that may be employed include purified PRO, fusion proteins containing PRO, and cells expressing recombinant PRO on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

10 Mice, such as Balb/c, are immunized with the PRO immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital  
15 bleeding for testing in ELISA assays to detect anti-PRO antibodies.

After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then  
20 be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells will be screened in an ELISA for reactivity against PRO. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against PRO is within the skill in the art.

25 The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

6.12. EXAMPLE 12: Purification of PRO Polypeptides Using Specific Antibodies

30 Native or recombinant PRO polypeptides may be purified by a variety of standard techniques in the art of protein purification. For example, pro-PRO polypeptide, mature PRO polypeptide, or pre-PRO polypeptide is purified by immunoaffinity chromatography using antibodies specific for the PRO polypeptide of interest. In general, an immunoaffinity column is constructed by covalently coupling the anti-PRO polypeptide antibody to an activated chromatographic resin.

35 Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.). Likewise,

monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated SEPHAROSE™ (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

Such an immunoaffinity column is utilized in the purification of PRO polypeptide by preparing a fraction from cells containing PRO polypeptide in a soluble form. This preparation is derived by solubilization of the whole cell or of a subcellular fraction obtained via differential centrifugation by the addition of detergent or by other methods well known in the art. Alternatively, soluble PRO polypeptide containing a signal sequence may be secreted in useful quantity into the medium in which the cells are grown.

A soluble PRO polypeptide-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PRO polypeptide (*e.g.*, high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/PRO polypeptide binding (*e.g.*, a low pH buffer such as approximately pH 2-3, or a high concentration of a chaotrope such as urea or thiocyanate ion), and PRO polypeptide is collected.

#### 6.13. EXAMPLE 13: Drug Screening

This invention is particularly useful for screening compounds by using PRO polypeptides or binding fragment thereof in any of a variety of drug screening techniques. The PRO polypeptide or fragment employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the PRO polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between PRO polypeptide or a fragment and the agent being tested. Alternatively, one can examine the diminution in complex formation between the PRO polypeptide and its target cell or target receptors caused by the agent being tested.

Thus, the present invention provides methods of screening for drugs or any other agents which can affect a PRO polypeptide-associated disease or disorder. These methods comprise contacting such an agent with an PRO polypeptide or fragment thereof and assaying (i) for the presence of a complex between the agent and the PRO polypeptide or fragment, or (ii) for the presence of a complex between the PRO polypeptide or fragment and the cell, by methods well known in the art. In such competitive binding assays, the PRO polypeptide or fragment is typically labeled. After suitable incubation, free PRO polypeptide or fragment is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular agent to bind to PRO polypeptide or to interfere with the PRO polypeptide/cell complex.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to a polypeptide and is described in detail in WO 84/03564, published on September 13, 1984. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such



as plastic pins or some other surface. As applied to a PRO polypeptide, the peptide test compounds are reacted with PRO polypeptide and washed. Bound PRO polypeptide is detected by methods well known in the art. Purified PRO polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the peptide and immobilize it on the solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding PRO polypeptide specifically compete with a test compound for binding to PRO polypeptide or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PRO polypeptide.

#### 6.14. EXAMPLE 14: Rational Drug Design

The goal of rational drug design is to produce structural analogs of biologically active polypeptide of interest (*i.e.*, a PRO polypeptide) or of small molecules with which they interact, *e.g.*, agonists, antagonists, or inhibitors. Any of these examples can be used to fashion drugs which are more active or stable forms of the PRO polypeptide or which enhance or interfere with the function of the PRO polypeptide *in vivo* (*c.f.*, Hodgson, Bio/Technology, 9: 19-21 (1991)).

In one approach, the three-dimensional structure of the PRO polypeptide, or of an PRO polypeptide-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the PRO polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of the PRO polypeptide may be gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design analogous PRO polypeptide-like molecules or to identify efficient inhibitors. Useful examples of rational drug design may include molecules which have improved activity or stability as shown by Braxton and Wells, Biochemistry, 31:7796-7801 (1992) or which act as inhibitors, agonists, or antagonists of native peptides as shown by Athauda *et al.*, J. Biochem., 113:742-746 (1993).

It is also possible to isolate a target-specific antibody, selected by functional assay, as described above, and then to solve its crystal structure. This approach, in principle, yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced peptides. The isolated peptides would then act as the pharmacore.

By virtue of the present invention, sufficient amounts of the PRO polypeptide may be made available to perform such analytical studies as X-ray crystallography. In addition, knowledge of the PRO polypeptide amino acid sequence provided herein will provide guidance to those employing computer modeling techniques in place of or in addition to x-ray crystallography.

6.15. EXAMPLE 15: Stimulation of Endothelial Cell Proliferation (Assay 8)

This assay is designed to determine whether PRO polypeptides of the present invention show the ability to stimulate adrenal cortical capillary endothelial cell (ACE) growth. PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of conditions or disorders where angiogenesis would be beneficial including, for example, wound healing, and the like (as would agonists of these PRO polypeptides). Antagonists of the PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of cancerous tumors.

Bovine adrenal cortical capillary endothelial (ACE) cells (from primary culture, maximum of 12-14 passages) were plated in 96-well plates at 500 cells/well per 100 microliter. Assay media included low glucose DMEM, 10% calf serum, 2 mM glutamine, and 1X penicillin/streptomycin/fungizone. Control wells included the following: (1) no ACE cells added; (2) ACE cells alone; (3) ACE cells plus VEGF (5 ng/ml); and (4) ACE cells plus FGF (5ng/ml). The control or test sample, (in 100 microliter volumes), was then added to the wells (at dilutions of 1%, 0.1% and 0.01%, respectively). The cell cultures were incubated for 6-7 days at 37°C/5% CO<sub>2</sub>. After the incubation, the media in the wells was aspirated, and the cells were washed 1X with PBS. An acid phosphatase reaction mixture (100 microliter; 0.1M sodium acetate, pH 5.5, 0.1% Triton X-100, 10 mM p-nitrophenyl phosphate) was then added to each well. After a 2 hour incubation at 37°C, the reaction was stopped by addition of 10 microliters 1N NaOH. Optical density (OD) was measured on a microplate reader at 405 nm.

The activity of a PRO polypeptide was calculated as the fold increase in proliferation (as determined by the acid phosphatase activity, OD 405 nm) relative to (1) cell only background, and (2) relative to maximum stimulation by VEGF. VEGF (at 3-10 ng/ml) and FGF (at 1-5 ng/ml) were employed as an activity reference for maximum stimulation. Results of the assay were considered "positive" if the observed stimulation was  $\geq$  50% increase over background. VEGF (5 ng/ml) control at 1% dilution gave 1.24 fold stimulation; FGF (5 ng/ml) control at 1% dilution gave 1.46 fold stimulation.

PRO21 tested positive in this assay.

6.16. EXAMPLE 16: Inhibition of Vascular Endothelial Growth Factor (VEGF) Stimulated Proliferation of Endothelial Cell Growth (Assay 9)

The ability of various PRO polypeptides to inhibit VEGF stimulated proliferation of endothelial cells was tested. Polypeptides testing positive in this assay are useful for inhibiting endothelial cell growth in mammals where such an effect would be beneficial, e.g., for inhibiting tumor growth.

Specifically, bovine adrenal cortical capillary endothelial cells (ACE) (from primary culture, maximum of 12-14 passages) were plated in 96-well plates at 500 cells/well per 100 microliter. Assay media included low glucose DMEM, 10% calf serum, 2 mM glutamine, and 1X penicillin/streptomycin/fungizone. Control wells included the following: (1) no ACE cells added; (2) ACE cells alone; (3) ACE cells plus 5 ng/ml FGF; (4) ACE cells plus 3 ng/ml VEGF; (5) ACE cells plus 3 ng/ml VEGF plus 1 ng/ml TGF-beta; and (6) ACE cells plus 3 ng/ml VEGF plus 5 ng/ml LIF. The test samples, poly-his tagged PRO polypeptides (in 100 microliter volumes), were then added to the wells (at dilutions of 1%, 0.1% and 0.01%, respectively). The cell cultures were

incubated for 6-7 days at 37°C/5% CO<sub>2</sub>. After the incubation, the media in the wells was aspirated, and the cells were washed 1X with PBS. An acid phosphatase reaction mixture (100 microliter; 0.1M sodium acetate, pH 5.5, 0.1% Triton X-100, 10 mM p-nitrophenyl phosphate) was then added to each well. After a 2 hour incubation at 37°C, the reaction was stopped by addition of 10 microliters 1N NaOH. Optical density (OD) was measured on a microplate reader at 405 nm.

The activity of PRO polypeptides was calculated as the percent inhibition of VEGF (3 ng/ml) stimulated proliferation (as determined by measuring acid phosphatase activity at OD 405 nm) relative to the cells without stimulation. TGF-beta was employed as an activity reference at 1 ng/ml, since TGF-beta blocks 70-90% of VEGF-stimulated ACE cell proliferation. The results are indicative of the utility of the PRO polypeptides in cancer therapy and specifically in inhibiting tumor angiogenesis. Numerical values (relative inhibition) are determined by calculating the percent inhibition of VEGF stimulated proliferation by the PRO polypeptides relative to cells without stimulation and then dividing that percentage into the percent inhibition obtained by TGF-β at 1 ng/ml which is known to block 70-90% of VEGF stimulated cell proliferation. The results are considered positive if the PRO polypeptide exhibits 30% or greater inhibition of VEGF stimulation of endothelial cell growth (relative inhibition 30% or greater).

PRO247, PRO720 and PRO4302 tested positive in this assay.

6.17. EXAMPLE 17: Enhancement of Heart Neonatal Hypertrophy Induced by LIF+ET-1 (Assay 75)

This assay is designed to determine whether PRO polypeptides of the present invention show the ability to enhance neonatal heart hypertrophy induced by LIF and endothelin-1 (ET-1). A test compound that provides a positive response in the present assay would be useful for the therapeutic treatment of cardiac insufficiency diseases or disorders characterized or associated with an undesired level of hypertrophy of the cardiac muscle.

Cardiac myocytes from 1-day old Harlan Sprague Dawley rats (180 μl at 7.5 x 10<sup>4</sup>/ml, serum < 0.1, freshly isolated) are introduced on day 1 to 96-well plates previously coated with DMEM/F12 + 4%FCS. Test PRO polypeptide samples or growth medium alone (negative control) are then added directly to the wells on day 2 in 20 μl volume. LIF + ET-1 are then added to the wells on day 3. The cells are stained after an additional 2 days in culture and are then scored visually the next day. A positive in the assay occurs when the PRO polypeptide treated myocytes obtain a score greater than zero. A score of zero represents non-responsive cells whereas scores of 1 or 2 represent enhancement (*i.e.* they are visually larger on the average or more numerous than the untreated myocytes).

PRO21 polypeptides tested positive in this assay.

6.18. EXAMPLE 18: Detection of Endothelial Cell Apoptosis (FACS) (Assay 96)

The ability of PRO polypeptides of the present invention to induce apoptosis in endothelial cells was tested in human venous umbilical vein endothelial cells (HUVEC, Cell Systems) in gelatinized T175 flasks using HUVEC cells below passage 10. PRO polypeptides testing positive in this assay are expected to be useful for

therapeutically treating conditions where apoptosis of endothelial cells would be beneficial including, for example, the therapeutic treatment of tumors.

On day one, the cells were split [420,000 cells per gelatinized 6 cm dishes - ( $11 \times 10^3$  cells/cm<sup>2</sup> Falcon, Primaria)] and grown in media containing serum (CS-C, Cell System) overnight or for 16 hours to 24 hours.

On day 2, the cells were washed 1x with 5 ml PBS ; 3 ml of 0% serum medium was added with VEGF (100 ng/ml); and 30  $\mu$ l of the PRO test compound (final dilution 1%) or 0% serum medium (negative control) was added. The mixtures were incubated for 48 hours before harvesting.

The cells were then harvested for FACS analysis. The medium was aspirated and the cells washed once with PBS. 5 ml of 1 x trypsin was added to the cells in a T-175 flask, and the cells were allowed to stand until they were released from the plate (about 5-10 minutes). Trypsinization was stopped by adding 5 ml of growth media. The cells were spun at 1000 rpm for 5 minutes at 4°C. The media was aspirated and the cells were resuspended in 10 ml of 10% serum complemented medium (Cell Systems), 5  $\mu$ l of Annexin-FITC (BioVison) added and chilled tubes were submitted for FACS. A positive result was determined to be enhanced apoptosis in the PRO polypeptide treated samples as compared to the negative control.

PRO4302 polypeptide tested positive in this assay.

#### 6.19. EXAMPLE 19: Induction of c-fos in HUVEC Cells (Assay 123)

This assay is designed to determine whether PRO polypeptides show the ability to induce c-fos in HUVEC cells. PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of conditions or disorders where angiogenesis would be beneficial including, for example, wound healing, and the like (as would agonists of these PRO polypeptides). Antagonists of the PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of cancerous tumors.

Human venous umbilical vein endothelial cells (HUVEC, Cell Systems) in growth media (50% Ham's F12 w/o GHT: low glucose, and 50% DMEM without glycine: with NaHCO<sub>3</sub>, 1% glutamine, 10 mM HEPES, 10% FBS, 10 ng/ml bFGF) were plated on 96-well microtiter plates at a cell density of  $5 \times 10^3$  cells/well. The day after plating (day 2), the cells were starved for 24 hours by removing the growth media and replacing with serum free media. On day 3, the cells are treated with 100  $\mu$ l/well test samples and controls (positive control = growth media; negative control = Protein 32 buffer = 10 mM HEPES, 140 mM NaCl, 4% (w/v) mannitol, pH 6.8). One plate of cells was incubated for 30 minutes at 37°C, in 5% CO<sub>2</sub>. Another plate of cells was incubated for 60 minutes at 37°C, in 5% CO<sub>2</sub>. The samples were removed, and RNA was harvested using the RNeasy 96 kit (Qiagen). Next, the RNA was assayed for c-fos, egr-1 and GAPDH induction using Taqman.

The measure of activity of the fold increase over the negative control (Protein 32/HEPES buffer described above) value was by obtained by calculating the fold increase of the ratio of c-fos to GAPDH in test samples as compared to the negative control. The results are considered positive if the PRO polypeptide exhibits at least a two-fold value over the negative buffer control.

PRO1376 polypeptide tested positive in this assay.

6.20. EXAMPLE 20: Normal Human Iliac Artery Endothelial Cell Proliferation (Assay 138)

This assay is designed to determine whether PRO polypeptides of the present invention show the ability to induce proliferation of human iliac artery endothelial cells in culture and, therefore, function as useful growth factors.

On day 0, human iliac artery endothelial cells (from cell lines, maximum of 12-14 passages) were plated in 96-well plates at 1000 cells/well per 100 microliter and incubated overnight in complete media [epithelial cell growth media (EGM, Clonetics), plus supplements: human epithelial growth factor (hEGF), bovine brain extract (BBE), hydrocortisone, GA-1000, and fetal bovine serum (FBS, Clonetics)]. On day 1, complete media was replaced by basal media [EGM plus 1% FBS] and addition of PRO polypeptides at 1%, 0.1% and 0.01%. On day 7, an assessment of cell proliferation was performed by Alamar Blue assay followed by Crystal Violet. Results are expressed as % of the cell growth observed with control buffer.

The following PRO polypeptides tested positive in this assay: PRO214, PRO238, PRO256, PRO363, PRO365, PRO791, PRO836, PRO1025, PRO1029, PRO1186, PRO1192, PRO1272, PRO1274, PRO1279, PRO1306, PRO1325, PRO1329, PRO1376, PRO1411, PRO1419, PRO1508, PRO1787, PRO1868, PRO1890, PRO4324, PRO4333, PRO4408, PRO4499, PRO5725, PRO6006, PRO9821, PRO9873, PRO10008, PRO10096, PRO19670, PRO20040, PRO20044, PRO21384 and PRO28631.

6.21. EXAMPLE 21: Pooled Human Umbilical Vein Endothelial Cell Proliferation (Assay 139)

This assay is designed to determine whether PRO polypeptides of the present invention show the ability to induce proliferation of pooled human umbilical vein endothelial cells in culture and, therefore, function as useful growth factors.

On day 0, pooled human umbilical vein endothelial cells (from cell lines, maximum of 12-14 passages) were plated in 96-well plates at 1000 cells/well per 100 microliter and incubated overnight in complete media [epithelial cell growth media (EGM, Clonetics), plus supplements: human epithelial growth factor (hEGF), bovine brain extract (BBE), hydrocortisone, GA-1000, and fetal bovine serum (FBS, Clonetics)]. On day 1, complete media was replaced by basal media [EGM plus 1% FBS] and addition of PRO polypeptides at 1%, 0.1% and 0.01%. On day 7, an assessment of cell proliferation was performed by Alamar Blue assay followed by Crystal Violet. Results are expressed as % of the cell growth observed with control buffer.

The following PRO polypeptides tested positive in this assay: PRO181, PRO205, PRO221, PRO229, PRO231, PRO238, PRO241, PRO247, PRO256, PRO258, PRO263, PRO265, PRO295, PRO321, PRO322, PRO337, PRO363, PRO444, PRO533, PRO697, PRO725, PRO771, PRO788, PRO819, PRO827, PRO828, PRO846, PRO865, PRO1005, PRO1006, PRO1007, PRO1025, PRO1054, PRO1071, PRO1075, PRO1079, PRO1080, PRO1114, PRO1131, PRO1155, PRO1160, PRO1184, PRO1190, PRO1192, PRO1195, PRO1244, PRO1272, PRO1273, PRO1279, PRO1283, PRO1286, PRO1306, PRO1309, PRO1325, PRO1329, PRO1347, PRO1356, PRO1376, PRO1382, PRO1412, PRO1419, PRO1474, PRO1477, PRO1488, PRO1550, PRO1556, PRO1760, PRO1782, PRO1787, PRO1801, PRO1868, PRO1887, PRO3438, PRO3444, PRO4302, PRO4324, PRO4341, PRO4342, PRO4353, PRO4354, PRO4356, PRO4371, PRO4405, PRO4422, PRO4425, PRO5723,

PRO5725, PRO5737, PRO5776, PRO6029, PRO6071, PRO7436, PRO9771, PRO10008, PRO10096, PRO21055 and PRO21384.

6.22. EXAMPLE 22: Human Coronary Artery Smooth Muscle Cell Proliferation (Assay 140)

5 This assay is designed to determine whether PRO polypeptides of the present invention show the ability to induce proliferation of human coronary artery smooth muscle cells in culture and, therefore, function as useful growth factors.

On day 0, human coronary artery smooth muscle cells (from cell lines, maximum of 12-14 passages) were plated in 96-well plates at 1000 cells/well per 100 microliter and incubated overnight in complete media [smooth muscle growth media (SmGM, Clonetics), plus supplements: insulin, human epithelial growth factor (hEGF), human fibroblast growth factor (hFGF), GA-1000, and fetal bovine serum (FBS, Clonetics)]. On day 10 1, complete media was replaced by basal media [SmGM plus 1% FBS] and addition of PRO polypeptides at 1%, 0.1% and 0.01%. On day 7, an assessment of cell proliferation was performed by Alamar Blue assay followed by Crystal Violet. Results are expressed as % of the cell growth observed with control buffer.

The following PRO polypeptides tested positive in this assay: PRO162, PRO181, PRO182, PRO195, 15 PRO204, PRO221, PRO230, PRO256, PRO258, PRO533, PRO697, PRO725, PRO738, PRO826, PRO836, PRO840, PRO846, PRO865, PRO982, PRO1025, PRO1029, PRO1071, PRO1080, PRO1083, PRO1134, PRO1160, PRO1182, PRO1184, PRO1186, PRO1192, PRO1265, PRO1274, PRO1279, PRO1283, PRO1306, PRO1308, PRO1309, PRO1325, PRO1337, PRO1338, PRO1343, PRO1376, PRO1387, PRO1411, PRO1412, PRO1415, PRO1434, PRO1474, PRO1488, PRO1550, PRO1556, PRO1567, PRO1600, PRO1754, PRO1758, PRO1760, 20 PRO1787, PRO1865, PRO1868, PRO1917, PRO1928, PRO3438, PRO3562, PRO4302, PRO4333, PRO4345, PRO4353, PRO4354, PRO4405, PRO4408, PRO4430, PRO4503, PRO5725, PRO6714, PRO9771, PRO9820, PRO9940, PRO10096, PRO21055, PRO21184 and PRO21366.

6.23. EXAMPLE 23: Microarray Analysis to Detect Overexpression of PRO Polypeptides in HUVEC Cells Treated with Growth Factors

25 This assay is designed to determine whether PRO polypeptides of the present invention show the ability to induce angiogenesis by stimulating endothelial cell tube formation in HUVEC cells.

Nucleic acid microarrays, often containing thousands of gene sequences, are useful for identifying differentially expressed genes in tissues exposed to various stimuli (*e.g.*, growth factors) as compared to their normal, unexposed counterparts. Using nucleic acid microarrays, test and control mRNA samples from test and control tissue samples are reverse transcribed and labeled to generate cDNA probes. The cDNA probes are then 30 hybridized to an array of nucleic acids immobilized on a solid support. The array is configured such that the sequence and position of each member of the array is known. Hybridization of a labeled probe with a particular array member indicates that the sample from which the probe was derived expresses that gene. If the hybridization signal of a probe from a test (exposed tissue) sample is greater than hybridization signal of a probe from a control (normal, unexposed tissue) sample, the gene or genes overexpressed in the exposed tissue are identified. The 35

implication of this result is that an overexpressed protein in an exposed tissue may be involved in the functional changes within the tissue following exposure to the stimuli (e.g., tube formation).

The methodology of hybridization of nucleic acids and microarray technology is well known in the art. In the present example, the specific preparation of nucleic acids for hybridization and probes, slides, and hybridization conditions are all detailed in U.S. Provisional Patent Application Serial No. 60/193,767, filed on March 31, 2000 and which is herein incorporated by reference.

In the present example, HUVEC cells grown in either collagen gels or fibrin gels were induced to form tubes by the addition of various growth factors. Specifically, collagen gels were prepared as described previously in Yang *et al.*, *American J. Pathology*, 1999, 155(3):887-895 and Xin *et al.*, *American J. Pathology*, 2001, 158(3):1111-1120. Following gelation of the HUVEC cells, 1X basal medium containing M199 supplemented with 1%FBS, 1X ITS, 2 mM L-glutamine, 50 µg/ml ascorbic acid, 26.5 mM NaHCO<sub>3</sub>, 100U/ml penicillin and 100 U/ml streptomycin was added. Tube formation was elicited by the inclusion in the culture media of either a mixture of phorbol myrsitate acetate (50 nM), vascular endothelial cell growth factor (40 ng/ml) and basic fibroblast growth factor (40 ng/ml) ("PMA growth factor mix") or hepatocyte growth factor (40 ng/ml) and vascular endothelial cell growth factor (40 ng/ml) (HGF/VEGF mix) for the indicated period of time. Fibrin Gels were prepared by suspending Huvec (4 x 10<sup>5</sup> cells/ml) in M199 containing 1% fetal bovine serum (Hyclone) and human fibrinogen (2.5mg/ml). Thrombin (50U/ml) was then added to the fibrinogen suspension at a ratio of 1 part thrombin solution:30 parts fibrinogen suspension. The solution was then layered onto 10 cm tissue culture plates (total volume: 15 ml/plate) and allowed to solidify at 37°C for 20 min. Tissue culture media (10 ml of BM containing PMA (50 nM), bFGF (40ng/ml) and VEGF (40 ng/ml)) was then added and the cells incubated at 37°C in 5%CO<sub>2</sub> in air for the indicated period of time.

Total RNA was extracted from the HUVEC cells incubated for 0, 4, 8, 24, 40 and 50 hours in the different matrix and media combinations using a TRIzol extraction followed by a second purification using RNAeasy Mini Kit (Qiagen). The total RNA was used to prepare cRNA which was then hybridized to the microarrays.

In the present experiments, nucleic acid probes derived from the herein described PRO polypeptide-encoding nucleic acid sequences were used in the creation of the microarray and RNA from the HUVEC cells described above were used for the hybridization thereto. Pairwise comparisons were made using time 0 chips as a baseline. Three replicate samples were analyzed for each experimental condition and time. Hence there were 3 time 0 samples for each treatment and 3 replicates of each successive time point. Therefore, a 3 by 3 comparison was performed for each time point compared against each time 0 point. This resulted in 9 comparisons per time point. Only those genes that had increased expression in all three non-time-0 replicates in each of the different matrix and media combinations as compared to any of the three time zero replicates were considered positive. Although this stringent method of data analysis does allow for false negatives, it minimizes false positives.

PRO178, PRO195, PRO228, PRO301, PRO302, PRO532, PRO724, PRO730, PRO734, PRO793, PRO871, PRO938, PRO1012, PRO1120, PRO1139, PRO1198, PRO1287, PRO1361, PRO1864, PRO1873,

PRO2010, PRO3579, PRO4313, PRO4527, PRO4538, PRO4553, PRO4995, PRO5730, PRO6008, PRO7223, PRO7248 and PRO7261 tested positive in this assay.

6.24. EXAMPLE 24: *In situ* Hybridization

*In situ* hybridization is a powerful and versatile technique for the detection and localization of nucleic acid sequences within cell or tissue preparations. It may be useful, for example, to identify sites of gene expression, analyze the tissue distribution of transcription, identify and localize viral infection, follow changes in specific mRNA synthesis, and aid in chromosome mapping.

*In situ* hybridization was performed following an optimized version of the protocol by Lu and Gillett, Cell Vision, 1: 169-176 (1994), using PCR-generated <sup>32</sup>P-labeled riboprobes. Briefly, formalin-fixed, paraffin-embedded human tissues were sectioned, deparaffinized, deproteinized in proteinase K (20 g/ml) for 15 minutes at 37°C, and further processed for *in situ* hybridization as described by Lu and Gillett, *supra*. A (<sup>32</sup>-P)UTP-labeled antisense riboprobe was generated from a PCR product and hybridized at 55°C overnight. The slides were dipped in Kodak NTB2™ nuclear track emulsion and exposed for 4 weeks.

6.24.1. <sup>32</sup>P-Riboprobe synthesis

6.0 µl (125 mCi) of <sup>32</sup>P-UTP (Amersham BF 1002, SA < 2000 Ci/mmol) were speed-vacuum dried. To each tube containing dried <sup>32</sup>P-UTP, the following ingredients were added:

2.0 µl 5x transcription buffer

1.0 µl DTT (100 mM)

2.0 µl NTP mix (2.5 mM: 10 µl each of 10 mM GTP, CTP & ATP + 10 µl H<sub>2</sub>O)

1.0 µl UTP (50 µM)

1.0 µl RNAsin

1.0 µl DNA template (1 µg)

1.0 µl H<sub>2</sub>O

1.0 µl RNA polymerase (for PCR products T3 = AS, T7 = S, usually)

The tubes were incubated at 37°C for one hour. A total of 1.0 µl RQ1 DNase was added, followed by incubation at 37°C for 15 minutes. A total of 90 µl TE (10 mM Tris pH 7.6/1 mM EDTA pH 8.0) was added, and the mixture was pipetted onto DE81 paper. The remaining solution was loaded in a MICROCON-50™ ultrafiltration unit, and spun using program 10 (6 minutes). The filtration unit was inverted over a second tube and spun using program 2 (3 minutes). After the final recovery spin, a total of 100 µl TE was added, then 1 µl of the final product was pipetted on DE81 paper and counted in 6 ml of BIOFLUOR II™.

The probe was run on a TBE/urea gel. A total of 1-3 µl of the probe or 5 µl of RNA Mrk III was added to 3 µl of loading buffer. After heating on a 95°C heat block for three minutes, the gel was immediately placed on ice. The wells of gel were flushed, and the sample was loaded and run at 180-250 volts for 45 minutes. The gel was wrapped in plastic wrap (SARAN™ brand) and exposed to XAR film with an intensifying screen in a -70°C freezer one hour to overnight.



6.24.2. <sup>32</sup>P-Hybridization6.24.2.1. *Pretreatment of frozen sections*

5 The slides were removed from the freezer, placed on aluminum trays, and thawed at room temperature for 5 minutes. The trays were placed in a 55°C incubator for five minutes to reduce condensation. The slides were fixed for 10 minutes in 4% paraformaldehyde on ice in the fume hood, and washed in 0.5 x SSC for 5 minutes, at room temperature (25 ml 20 x SSC + 975 ml SQ H<sub>2</sub>O). After deproteinization in 0.5 µg/ml proteinase K for 10 minutes at 37°C (12.5 µl of 10 mg/ml stock in 250 ml prewarmed RNase-free RNase buffer), the sections were washed in 0.5 x SSC for 10 minutes at room temperature. The sections were dehydrated in 70%, 95%, and 100% ethanol, 2 minutes each.

6.24.2.2. *Pretreatment of paraffin-embedded sections*

10 The slides were deparaffinized, placed in SQ H<sub>2</sub>O, and rinsed twice in 2 x SSC at room temperature, for 5 minutes each time. The sections were deproteinized in 20 µg/ml proteinase K (500 µl of 10 mg/ml in 250 ml RNase-free RNase buffer; 37°C, 15 minutes) for human embryo tissue, or 8 x proteinase K (100 µl in 250 ml RNase buffer, 37°C, 30 minutes) for formalin tissues. Subsequent rinsing in 0.5 x SSC and dehydration were performed as described above.

6.24.2.3. *Prehybridization*

20 The slides were laid out in a plastic box lined with Box buffer (4 x SSC, 50% formamide) - saturated filter paper. The tissue was covered with 50 µl of hybridization buffer (3.75 g dextran sulfate + 6 ml SQ H<sub>2</sub>O), vortexed, and heated in the microwave for 2 minutes with the cap loosened. After cooling on ice, 18.75 ml formamide, 3.75 ml 20 x SSC, and 9 ml SQ H<sub>2</sub>O were added, and the tissue was vortexed well and incubated at 42°C for 1-4 hours.

6.24.2.4. *Hybridization*

25 1.0 x 10<sup>6</sup> cpm probe and 1.0 µl tRNA (50 mg/ml stock) per slide were heated at 95°C for 3 minutes. The slides were cooled on ice, and 48 µl hybridization buffer was added per slide. After vortexing, 50 µl <sup>32</sup>P mix was added to 50 µl prehybridization on the slide. The slides were incubated overnight at 55°C.

6.24.2.5. *Washes*

30 Washing was done for 2x10 minutes with 2xSSC, EDTA at room temperature (400 ml 20 x SSC + 16 ml 0.25 M EDTA, V<sub>f</sub>=4L), followed by RNaseA treatment at 37°C for 30 minutes (500 µl of 10 mg/ml in 250 ml RNase buffer = 20 µg/ml). The slides were washed 2 x10 minutes with 2 x SSC, EDTA at room temperature. The stringency wash conditions were as follows: 2 hours at 55°C, 0.1 x SSC, EDTA (20 ml 20 x SSC + 16 ml EDTA, V<sub>f</sub>=4L).

6.24.2.6. *Oligonucleotides*

*In situ* analysis was performed on three of the DNA sequences disclosed herein. The primers used to generate the probes and/or the probes employed for these analyses are as follows:

- 5 DNA33100-p1: 5'GGA TTC TAA TAC GAC TCA CTA TAG GGC CGG GTG GAG GTG GAA CAG AAA  
3' (SEQ ID NO:375)
- DNA33100-p2: 5' CTA TGA AAT TAA CCC TCA CTA AAG GGA CAC AGA CAG AGC CCC ATA CGC  
3' (SEQ ID NO:376)
- DNA34431-p1: 5'GGA TTC TAA TAC GAC TCA CTA TAG GGC CAG GGA AAT CCG GAT GTC TC  
3' (SEQ ID NO:377)
- 10 DNA34431-p2: 5' CTA TGA AAT TAA CCC TCA CTA AAG GGA GTA AGG GGA TGC CAC CGA GTA  
3' (SEQ ID NO:378)
- DNA38268-p1: 5'GGA TTC TAA TAC GAC TCA CTA TAG GGC CAG CTA CCC GCA GGA GGA GG  
3' (SEQ ID NO:379)
- DNA38268-p2: 5'CTA TGA AAT TAA CCC TCA CTA AAG GGA TCC CAG GTG ATG AGG TCC AGA  
15 3' (SEQ ID NO:380)
- DNA64908 probe: 5'CCATCTCGGAGACCTTTGTGCAGCGTGTATACCAGCCTTACCTCACCA  
CTTGCGACGGACACAGAGCCTGCAGCACCTACCGAACCATCTACCGGAC  
TGCCCTATCGCCGTAGCCCTGGGGTGACTCCCGCAAGCCTCGCTATGCTTG  
CTGCCCTGGTTGGAAGAGGACCAGTGGGCTCCCTGGGGCTTGTGGAGCA  
20 GCAATATGCCAGCCTCCATGTGGGAATGGAGGGAGTTGCATCCGCCCAG  
GACACTGCCGCTGCCCTGTGGGATGGCAGGGAGATACTTGCCAGACAGA  
TGTTGATGAATGCAGTACAGGAGAGGCCAGTTGTCCCCAGCGCTGTGTC  
AATACTGTGGGAAGTTACTGGTGCCAGGGATGGGAGGGACAAAGCCCAT  
CTGCAGATGGGACGCGCTGCCTGTCTAAGGAGGGGCCCTCCCGGTGGCC  
25 CCAACCCACAGCAGGAGTGGACAGCA3' (SEQ ID NO:381)

6.24.2.7. *Results*

*In situ* analysis was performed and the results from these analyses are as follows:

6.24.2.7.1. DNA33100-1159 (PRO229) (Scavenger-R/CD6  
homologTNF motif)

A specific positive signal was observed in mononuclear phagocytes (macrophages) of fetal and adult spleen, liver, lymph node and thymus. All other tissues were negative.

5 6.24.2.7.2. DNA34431-1177 (PRO263) (CD44)

A specific positive signal was observed in human fetal tissues and placenta over mononuclear cells, with strong expression in epithelial cells of the adrenal cortex. All adult tissues were negative.

6.24.2.7.3. DNA38268-1188 (PRO295) (Integrin)

10 A specific positive signal was observed in human fetal ganglion cells, fetal neurons, adult adrenal medulla and adult neurons. All other tissues were negative.

6.24.2.7.4. DNA64908-1163-1 (PRO1449)

15 A specific positive signal was observed in the developing vasculature (from E7-E11), in endothelial cells and in progenitors of endothelial cells in wholemount in situ hybridizations of mouse embryos (Figure 375). Specific expression was also observed in a subset of blood vessels and epidermis from E12 onward. A mouse orthologue of PRO1449 which has about 78 % amino acid identity with PRO1449 was used as the probe.

In normal adult tissues, expression was low to absent. When present, expression was confined to the vasculature (Figure 376). Figure 376 further shows that highest expression in adult tissues was observed regionally in vessels running within the white matter of the brain. Elevated expression was also observed in vasculature of many inflamed and diseased tissues, including, but not limited to, tumor vasculature.

20 Following electroporation of the mouse orthologue of PRO1449 into the choroid layer in the eyes of chicken embryos, new vessel formation was observed in the electroporated eye (top right), but not in the control side from the same embryo (top left), or an embryo that was electroporated with a control cDNA (bottom right) (Figure 377).

25 The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct(s) deposited, since the deposited embodiment(s) is/are intended as single illustration(s) of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material(s) herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific  
30 illustrations that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule having at least 80% nucleic acid sequence identity to a nucleotide sequence that encodes an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure

248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372) and Figure 374 (SEQ ID NO:374).

2. An isolated nucleic acid molecule having at least 80% nucleic acid sequence identity to a nucleotide sequence selected from the group consisting of the nucleotide sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), Figure 5 (SEQ ID NO:5), Figure 7 (SEQ ID NO:7), Figure 9 (SEQ ID NO:9), Figure 11 (SEQ ID NO:11), Figure 13 (SEQ ID NO:13), Figure 15 (SEQ ID NO:15), Figure 17 (SEQ ID NO:17), Figure 19 (SEQ ID NO:19), Figure 21 (SEQ ID NO:21), Figure 23 (SEQ ID NO:23), Figure 25 (SEQ ID NO:25), Figure 27 (SEQ ID NO:27), Figure 29 (SEQ ID NO:29), Figure 31 (SEQ ID NO:31), Figure 33 (SEQ ID NO:33), Figure 35 (SEQ ID NO:35), Figure 37 (SEQ ID NO:37), Figure 39 (SEQ ID NO:39), Figure 41 (SEQ ID NO:41), Figure 43 (SEQ ID NO:43), Figure 45 (SEQ ID NO:45), Figure 47 (SEQ ID NO:47), Figure 49 (SEQ ID NO:49), Figure 51 (SEQ ID NO:51), Figure 53 (SEQ ID NO:53), Figure 55 (SEQ ID NO:55), Figure 57 (SEQ ID NO:57), Figure 59 (SEQ ID NO:59), Figure 61 (SEQ ID NO:61), Figure 63 (SEQ ID NO:63), Figure 65 (SEQ ID NO:65), Figure 67 (SEQ ID NO:67), Figure 69 (SEQ ID NO:69), Figure 71 (SEQ ID NO:71), Figure 73 (SEQ ID NO:73), Figures 75A-75B (SEQ ID NO:75), Figure 77 (SEQ ID NO:77), Figure 79 (SEQ ID NO:79), Figure 81 (SEQ ID NO:81), Figure 83 (SEQ ID NO:83), Figure 85 (SEQ ID NO:85), Figure 87 (SEQ ID NO:87), Figure 89 (SEQ ID NO:89), Figure 91 (SEQ ID NO:91), Figure 93 (SEQ ID NO:93), Figure 95 (SEQ ID NO:95), Figure 97 (SEQ ID NO:97), Figure 99 (SEQ ID NO:99), Figure 101 (SEQ ID NO:101), Figure 103 (SEQ ID NO:103), Figure 105 (SEQ ID NO:105), Figure 107 (SEQ ID NO:107), Figure 109 (SEQ ID NO:109), Figure 111 (SEQ ID NO:111), Figure 113 (SEQ ID NO:113), Figure 115 (SEQ

5 ID NO:115), Figure 117 (SEQ ID NO:117), Figure 119 (SEQ ID NO:119), Figure 121 (SEQ ID NO:121), Figure  
123 (SEQ ID NO:123), Figure 125 (SEQ ID NO:125), Figure 127 (SEQ ID NO:127), Figure 129 (SEQ ID  
NO:129), Figure 131 (SEQ ID NO:131), Figure 133 (SEQ ID NO:133), Figure 135 (SEQ ID NO:135), Figure  
137 (SEQ ID NO:137), Figure 139 (SEQ ID NO:139), Figure 141 (SEQ ID NO:141), Figure 143 (SEQ ID  
NO:143), Figure 145 (SEQ ID NO:145), Figure 147 (SEQ ID NO:147), Figure 149 (SEQ ID NO:149), Figure  
151 (SEQ ID NO:151), Figure 153 (SEQ ID NO:153), Figure 155 (SEQ ID NO:155), Figure 157 (SEQ ID  
NO:157), Figure 159 (SEQ ID NO:159), Figure 161 (SEQ ID NO:161), Figure 163 (SEQ ID NO:163), Figure  
165 (SEQ ID NO:165), Figure 167 (SEQ ID NO:167), Figure 169 (SEQ ID NO:169), Figure 171 (SEQ ID  
NO:171), Figure 173 (SEQ ID NO:173), Figure 175 (SEQ ID NO:175), Figure 177 (SEQ ID NO:177), Figure  
10 179 (SEQ ID NO:179), Figure 181 (SEQ ID NO:181), Figure 183 (SEQ ID NO:183), Figure 185 (SEQ ID  
NO:185), Figure 187 (SEQ ID NO:187), Figure 189 (SEQ ID NO:189), Figure 191 (SEQ ID NO:191), Figure  
193 (SEQ ID NO:193), Figure 195 (SEQ ID NO:195), Figure 197 (SEQ ID NO:197), Figure 199 (SEQ ID  
NO:199), Figure 201 (SEQ ID NO:201), Figure 203 (SEQ ID NO:203), Figure 205 (SEQ ID NO:205), Figure  
207 (SEQ ID NO:207), Figure 209 (SEQ ID NO:209), Figure 211 (SEQ ID NO:211), Figure 213 (SEQ ID  
15 NO:213), Figure 215 (SEQ ID NO:215), Figure 217 (SEQ ID NO:217), Figure 219 (SEQ ID NO:219), Figure  
221 (SEQ ID NO:221), Figure 223 (SEQ ID NO:223), Figure 225 (SEQ ID NO:225), Figure 227 (SEQ ID  
NO:227), Figure 229 (SEQ ID NO:229), Figure 231 (SEQ ID NO:231), Figure 233 (SEQ ID NO:233), Figure  
235 (SEQ ID NO:235), Figure 237 (SEQ ID NO:237), Figure 239 (SEQ ID NO:239), Figure 241 (SEQ ID  
NO:241), Figure 243 (SEQ ID NO:243), Figure 245 (SEQ ID NO:245), Figure 247 (SEQ ID NO:247), Figure  
20 249 (SEQ ID NO:249), Figure 251 (SEQ ID NO:251), Figure 253 (SEQ ID NO:253), Figure 255 (SEQ ID  
NO:255), Figure 257 (SEQ ID NO:257), Figure 259 (SEQ ID NO:259), Figure 261 (SEQ ID NO:261), Figure  
263 (SEQ ID NO:263), Figure 265 (SEQ ID NO:265), Figure 267 (SEQ ID NO:267), Figure 269 (SEQ ID  
NO:269), Figure 271 (SEQ ID NO:271), Figure 273 (SEQ ID NO:273), Figure 275 (SEQ ID NO:275), Figure  
277 (SEQ ID NO:277), Figure 279 (SEQ ID NO:279), Figure 281 (SEQ ID NO:281), Figure 283 (SEQ ID  
25 NO:283), Figure 285 (SEQ ID NO:285), Figure 287 (SEQ ID NO:287), Figures 289A-289B (SEQ ID NO:289),  
Figure 291 (SEQ ID NO:291), Figure 293 (SEQ ID NO:293), Figure 295 (SEQ ID NO:295), Figure 297 (SEQ  
ID NO:297), Figure 299 (SEQ ID NO:299), Figure 301 (SEQ ID NO:301), Figure 303 (SEQ ID NO:303), Figure  
305 (SEQ ID NO:305), Figure 307 (SEQ ID NO:307), Figure 309 (SEQ ID NO:309), Figures 311A-311B (SEQ  
ID NO:311), Figure 313 (SEQ ID NO:313), Figure 315 (SEQ ID NO:315), Figure 317 (SEQ ID NO:317), Figure  
30 319 (SEQ ID NO:319), Figure 321 (SEQ ID NO:321), Figure 323 (SEQ ID NO:323), Figure 325 (SEQ ID  
NO:325), Figure 327 (SEQ ID NO:327), Figure 329 (SEQ ID NO:329), Figure 331 (SEQ ID NO:331), Figure  
333 (SEQ ID NO:333), Figure 335 (SEQ ID NO:335), Figure 337 (SEQ ID NO:337), Figure 339 (SEQ ID  
NO:339), Figure 341 (SEQ ID NO:341), Figure 343 (SEQ ID NO:343), Figure 345 (SEQ ID NO:345), Figure  
347 (SEQ ID NO:347), Figure 349 (SEQ ID NO:349), Figures 351A-351B (SEQ ID NO:351), Figure 353 (SEQ  
35 ID NO:353), Figure 355 (SEQ ID NO:355), Figure 357 (SEQ ID NO:357), Figure 359 (SEQ ID NO:359), Figure  
361 (SEQ ID NO:361), Figure 363 (SEQ ID NO:363), Figure 365 (SEQ ID NO:365), Figure 367 (SEQ ID  
NO:367), Figure 369 (SEQ ID NO:369), Figure 371 (SEQ ID NO:371) and Figure 373 (SEQ ID NO:373).

3. An isolated nucleic acid molecule having at least 80% nucleic acid sequence identity to a nucleotide sequence selected from the group consisting of the full-length coding sequence of the nucleotide sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), Figure 5 (SEQ ID NO:5), Figure 7 (SEQ ID NO:7), Figure 9 (SEQ ID NO:9), Figure 11 (SEQ ID NO:11), Figure 13 (SEQ ID NO:13), Figure 15 (SEQ ID NO:15), Figure 17 (SEQ ID NO:17), Figure 19 (SEQ ID NO:19), Figure 21 (SEQ ID NO:21), Figure 23 (SEQ ID NO:23), Figure 25 (SEQ ID NO:25), Figure 27 (SEQ ID NO:27), Figure 29 (SEQ ID NO:29), Figure 31 (SEQ ID NO:31), Figure 33 (SEQ ID NO:33), Figure 35 (SEQ ID NO:35), Figure 37 (SEQ ID NO:37), Figure 39 (SEQ ID NO:39), Figure 41 (SEQ ID NO:41), Figure 43 (SEQ ID NO:43), Figure 45 (SEQ ID NO:45), Figure 47 (SEQ ID NO:47), Figure 49 (SEQ ID NO:49), Figure 51 (SEQ ID NO:51), Figure 53 (SEQ ID NO:53), Figure 55 (SEQ ID NO:55), Figure 57 (SEQ ID NO:57), Figure 59 (SEQ ID NO:59), Figure 61 (SEQ ID NO:61), Figure 63 (SEQ ID NO:63), Figure 65 (SEQ ID NO:65), Figure 67 (SEQ ID NO:67), Figure 69 (SEQ ID NO:69), Figure 71 (SEQ ID NO:71), Figure 73 (SEQ ID NO:73), Figures 75A-75B (SEQ ID NO:75), Figure 77 (SEQ ID NO:77), Figure 79 (SEQ ID NO:79), Figure 81 (SEQ ID NO:81), Figure 83 (SEQ ID NO:83), Figure 85 (SEQ ID NO:85), Figure 87 (SEQ ID NO:87), Figure 89 (SEQ ID NO:89), Figure 91 (SEQ ID NO:91), Figure 93 (SEQ ID NO:93), Figure 95 (SEQ ID NO:95), Figure 97 (SEQ ID NO:97), Figure 99 (SEQ ID NO:99), Figure 101 (SEQ ID NO:101), Figure 103 (SEQ ID NO:103), Figure 105 (SEQ ID NO:105), Figure 107 (SEQ ID NO:107), Figure 109 (SEQ ID NO:109), Figure 111 (SEQ ID NO:111), Figure 113 (SEQ ID NO:113), Figure 115 (SEQ ID NO:115), Figure 117 (SEQ ID NO:117), Figure 119 (SEQ ID NO:119), Figure 121 (SEQ ID NO:121), Figure 123 (SEQ ID NO:123), Figure 125 (SEQ ID NO:125), Figure 127 (SEQ ID NO:127), Figure 129 (SEQ ID NO:129), Figure 131 (SEQ ID NO:131), Figure 133 (SEQ ID NO:133), Figure 135 (SEQ ID NO:135), Figure 137 (SEQ ID NO:137), Figure 139 (SEQ ID NO:139), Figure 141 (SEQ ID NO:141), Figure 143 (SEQ ID NO:143), Figure 145 (SEQ ID NO:145), Figure 147 (SEQ ID NO:147), Figure 149 (SEQ ID NO:149), Figure 151 (SEQ ID NO:151), Figure 153 (SEQ ID NO:153), Figure 155 (SEQ ID NO:155), Figure 157 (SEQ ID NO:157), Figure 159 (SEQ ID NO:159), Figure 161 (SEQ ID NO:161), Figure 163 (SEQ ID NO:163), Figure 165 (SEQ ID NO:165), Figure 167 (SEQ ID NO:167), Figure 169 (SEQ ID NO:169), Figure 171 (SEQ ID NO:171), Figure 173 (SEQ ID NO:173), Figure 175 (SEQ ID NO:175), Figure 177 (SEQ ID NO:177), Figure 179 (SEQ ID NO:179), Figure 181 (SEQ ID NO:181), Figure 183 (SEQ ID NO:183), Figure 185 (SEQ ID NO:185), Figure 187 (SEQ ID NO:187), Figure 189 (SEQ ID NO:189), Figure 191 (SEQ ID NO:191), Figure 193 (SEQ ID NO:193), Figure 195 (SEQ ID NO:195), Figure 197 (SEQ ID NO:197), Figure 199 (SEQ ID NO:199), Figure 201 (SEQ ID NO:201), Figure 203 (SEQ ID NO:203), Figure 205 (SEQ ID NO:205), Figure 207 (SEQ ID NO:207), Figure 209 (SEQ ID NO:209), Figure 211 (SEQ ID NO:211), Figure 213 (SEQ ID NO:213), Figure 215 (SEQ ID NO:215), Figure 217 (SEQ ID NO:217), Figure 219 (SEQ ID NO:219), Figure 221 (SEQ ID NO:221), Figure 223 (SEQ ID NO:223), Figure 225 (SEQ ID NO:225), Figure 227 (SEQ ID NO:227), Figure 229 (SEQ ID NO:229), Figure 231 (SEQ ID NO:231), Figure 233 (SEQ ID NO:233), Figure 235 (SEQ ID NO:235), Figure 237 (SEQ ID NO:237), Figure 239 (SEQ ID NO:239), Figure 241 (SEQ ID NO:241), Figure 243 (SEQ ID NO:243), Figure 245 (SEQ ID NO:245), Figure 247 (SEQ ID NO:247), Figure 249 (SEQ ID NO:249), Figure 251 (SEQ ID NO:251), Figure

253 (SEQ ID NO:253), Figure 255 (SEQ ID NO:255), Figure 257 (SEQ ID NO:257), Figure 259 (SEQ ID NO:259), Figure 261 (SEQ ID NO:261), Figure 263 (SEQ ID NO:263), Figure 265 (SEQ ID NO:265), Figure 267 (SEQ ID NO:267), Figure 269 (SEQ ID NO:269), Figure 271 (SEQ ID NO:271), Figure 273 (SEQ ID NO:273), Figure 275 (SEQ ID NO:275), Figure 277 (SEQ ID NO:277), Figure 279 (SEQ ID NO:279), Figure 281 (SEQ ID NO:281), Figure 283 (SEQ ID NO:283), Figure 285 (SEQ ID NO:285), Figure 287 (SEQ ID NO:287), Figures 289A-289B (SEQ ID NO:289), Figure 291 (SEQ ID NO:291), Figure 293 (SEQ ID NO:293), Figure 295 (SEQ ID NO:295), Figure 297 (SEQ ID NO:297), Figure 299 (SEQ ID NO:299), Figure 301 (SEQ ID NO:301), Figure 303 (SEQ ID NO:303), Figure 305 (SEQ ID NO:305), Figure 307 (SEQ ID NO:307), Figure 309 (SEQ ID NO:309), Figures 311A-311B (SEQ ID NO:311), Figure 313 (SEQ ID NO:313), Figure 315 (SEQ ID NO:315), Figure 317 (SEQ ID NO:317), Figure 319 (SEQ ID NO:319), Figure 321 (SEQ ID NO:321), Figure 323 (SEQ ID NO:323), Figure 325 (SEQ ID NO:325), Figure 327 (SEQ ID NO:327), Figure 329 (SEQ ID NO:329), Figure 331 (SEQ ID NO:331), Figure 333 (SEQ ID NO:333), Figure 335 (SEQ ID NO:335), Figure 337 (SEQ ID NO:337), Figure 339 (SEQ ID NO:339), Figure 341 (SEQ ID NO:341), Figure 343 (SEQ ID NO:343), Figure 345 (SEQ ID NO:345), Figure 347 (SEQ ID NO:347), Figure 349 (SEQ ID NO:349), Figures 351A-351B (SEQ ID NO:351), Figure 353 (SEQ ID NO:353), Figure 355 (SEQ ID NO:355), Figure 357 (SEQ ID NO:357), Figure 359 (SEQ ID NO:359), Figure 361 (SEQ ID NO:361), Figure 363 (SEQ ID NO:363), Figure 365 (SEQ ID NO:365), Figure 367 (SEQ ID NO:367), Figure 369 (SEQ ID NO:369), Figure 371 (SEQ ID NO:371) and Figure 373 (SEQ ID NO:373).

4. An isolated nucleic acid molecule having at least 80% nucleic acid sequence identity to the full-length coding sequence of the DNA deposited under any ATCC accession number shown in Table 7.

5. A vector comprising the nucleic acid of Claim 1.

6. A host cell comprising the vector of Claim 5.

7. The host cell of Claim 6, wherein said cell is a CHO cell.

8. The host cell of Claim 6, wherein said cell is an *E. coli*.

9. The host cell of Claim 6, wherein said cell is a yeast cell.

10. A process for producing a PRO polypeptide comprising culturing the host cell of Claim 6 under conditions suitable for expression of said PRO polypeptide and recovering said PRO polypeptide from the cell culture.



11. An isolated polypeptide having at least 80% amino acid sequence identity to an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID

NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372) and Figure 374 (SEQ ID NO:374).

12. An isolated polypeptide having at least 80% amino acid sequence identity to an amino acid sequence encoded by the full-length coding sequence of the DNA deposited under any ATCC accession number shown in Table 7.

13. A chimeric molecule comprising a polypeptide according to Claim 11 fused to a heterologous amino acid sequence.

14. The chimeric molecule of Claim 13, wherein said heterologous amino acid sequence is an epitope tag sequence.

15. The chimeric molecule of Claim 13, wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

16. An antibody which specifically binds to a polypeptide according to Claim 11.

17. The antibody of Claim 16, wherein said antibody is a monoclonal antibody, a humanized antibody or a single-chain antibody.

18. An isolated nucleic acid molecule having at least 80% nucleic acid sequence identity to:

(a) a nucleotide sequence encoding the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26),  
Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64),  
Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102),  
Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264).

NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372) or Figure 374 (SEQ ID NO:374), lacking its associated signal peptide;

(b) a nucleotide sequence encoding an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure

144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372) or Figure 374 (SEQ ID NO:374), with its associated signal peptide; or

(c) a nucleotide sequence encoding an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18

(SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure

284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372) or Figure 374 (SEQ ID NO:374), lacking its associated signal peptide.

19. An isolated polypeptide having at least 80% amino acid sequence identity to:

(a) an amino acid sequence of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID

NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372) or Figure 374 (SEQ ID NO:374), lacking its associated signal peptide;

(b) an amino acid sequence of an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26



(SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure

292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372) or Figure 374 (SEQ ID NO:374), with its associated signal peptide; or

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID

NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure  
 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID  
 NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure  
 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID  
 5 NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure  
 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID  
 NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure  
 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID  
 NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure  
 10 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID  
 NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure  
 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID  
 NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure  
 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID  
 15 NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure  
 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID  
 NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure  
 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID  
 NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure  
 20 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID  
 NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure  
 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID  
 NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure  
 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID  
 25 NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure  
 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID  
 NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure  
 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID  
 NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372) or Figure 374 (SEQ ID NO:374), lacking  
 30 its associated signal peptide.

20. A method for treating a cardiovascular, endothelial or angiogenic disorder in a mammal  
 comprising administering to the mammal a therapeutically effective amount of a polypeptide shown in Figure 2  
 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ  
 35 ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18  
 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure  
 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32),

Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure

298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372) or Figure 374 (SEQ ID NO:374), or agonist or antagonist thereof.

21. The method according to Claim 20, wherein the mammal is human.

15

22. The method of Claim 21, wherein the human has suffered myocardial infarction.

23. The method of Claim 21, wherein the human has cardiac hypertrophy, trauma, a cancer, or age-related macular degeneration.

20

24. The method of Claim 23, wherein the cardiac hypertrophy is characterized by the presence of an elevated level of  $\text{PGF}_{2\alpha}$ .

25. The method of Claim 20, wherein the polypeptide is administered together with a cardiovascular, endothelial or angiogenic agent.

25

26. The method of Claim 23, wherein the polypeptide is administered following primary angioplasty.

27. The method of Claim 20, wherein the cardiovascular, endothelial or angiogenic disorder is cancer.

30

28. The method of Claim 27, wherein the polypeptide is administered in combination with a chemotherapeutic agent, a growth inhibitory agent or a cytotoxic agent.

29. The method of Claim 20, wherein said agonist is an antibody to said polypeptide.

35

30. The method of Claim 20, wherein said antagonist is an antibody to said polypeptide.

31. A method for treating a cardiovascular, endothelial or angiogenic disorder in a mammal comprising administering to the mammal a nucleic acid molecule that encodes a polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure

256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372) or Figure 374 (SEQ ID NO:374), or agonist or antagonist thereof.

- 20           32.     The method of Claim 31, wherein said agonist is an antibody to said polypeptide.
33.     The method of Claim 31, wherein said antagonist is an antibody to said polypeptide.
34.     The method of Claim 31, wherein the mammal is human.
- 25           35.     The method of Claim 31, wherein the nucleic acid molecule is administered via *ex vivo* gene therapy.
36.     A method for inhibiting endothelial cell growth in a mammal comprising administering to the mammal a PRO247, PRO720 or PRO4302 polypeptide or agonist thereof, wherein endothelial cell growth in said mammal is inhibited.
- 30           37.     A method for stimulating endothelial cell growth in a mammal comprising administering to the mammal a PRO21, PRO181, PRO205, PRO214, PRO221, PRO229, PRO231, PRO238, PRO241, PRO247, PRO256, PRO258, PRO263, PRO265, PRO295, PRO321, PRO322, PRO337, PRO363, PRO365, PRO444, PRO533, PRO697, PRO725, PRO771, PRO788, PRO791, PRO819, PRO827, PRO828, PRO836, PRO846, PRO865, PRO1005, PRO1006, PRO1007, PRO1025, PRO1029, PRO1054, PRO1071, PRO1075, PRO1079,
- 35

PRO1080, PRO1114, PRO1131, PRO1155, PRO1160, PRO1184, PRO1186, PRO1190, PRO1192, PRO1195, PRO1244, PRO1272, PRO1273, PRO1274, PRO1279, PRO1283, PRO1286, PRO1306, PRO1309, PRO1325, PRO1329, PRO1347, PRO1356, PRO1376, PRO1382, PRO1411, PRO1412, PRO1419, PRO1474, PRO1477, PRO1488, PRO1508, PRO1550, PRO1556, PRO1760, PRO1782, PRO1787, PRO1801, PRO1868, PRO1887, PRO1890, PRO3438, PRO3444, PRO4302, PRO4324, PRO4333, PRO4341, PRO4342, PRO4353, PRO4354, PRO4356, PRO4371, PRO4405, PRO4408, PRO4422, PRO4425, PRO4499, PRO5723, PRO5725, PRO5737, PRO5776, PRO6006, PRO6029, PRO6071, PRO7436, PRO9771, PRO9821, PRO9873, PRO10008, PRO10096, PRO19670, PRO20040, PRO20044, PRO21055, PRO21384 or PRO28631 polypeptide, or agonist thereof, wherein endothelial cell growth in said mammal is stimulated.

38. A method for inducing cardiac hypertrophy in a mammal comprising administering to the mammal a PRO21 polypeptide or agonist thereof, wherein cardiac hypertrophy in said mammal is induced.

39. A method for stimulating angiogenesis induced by a PRO1376 or PRO1449 polypeptide in a mammal comprising administering a therapeutically effective amount of said polypeptide to the mammal, wherein said angiogenesis is stimulated.

40. A method for inducing endothelial cell apoptosis comprising administering to the endothelial cell a PRO4302 polypeptide or agonist thereof, wherein apoptosis in said endothelial cell is induced.

41. A method for stimulating smooth muscle cell growth comprising administering to the smooth muscle cell a PRO162, PRO181, PRO182, PRO195, PRO204, PRO221, PRO230, PRO256, PRO258, PRO533, PRO697, PRO725, PRO738, PRO826, PRO836, PRO840, PRO846, PRO865, PRO982, PRO1025, PRO1029, PRO1071, PRO1080, PRO1083, PRO1134, PRO1160, PRO1182, PRO1184, PRO1186, PRO1192, PRO1265, PRO1274, PRO1279, PRO1283, PRO1306, PRO1308, PRO1309, PRO1325, PRO1337, PRO1338, PRO1343, PRO1376, PRO1387, PRO1411, PRO1412, PRO1415, PRO1434, PRO1474, PRO1488, PRO1550, PRO1556, PRO1567, PRO1600, PRO1754, PRO1758, PRO1760, PRO1787, PRO1865, PRO1868, PRO1917, PRO1928, PRO3438, PRO3562, PRO4302, PRO4333, PRO4345, PRO4353, PRO4354, PRO4405, PRO4408, PRO4430, PRO4503, PRO5725, PRO6714, PRO9771, PRO9820, PRO9940, PRO10096, PRO21055, PRO21184 or PRO21366 polypeptide, or agonist thereof, wherein smooth muscle cell growth in said smooth muscle cell is stimulated.

42. A method for inducing endothelial cell tube formation comprising administering to the endothelial cell a PRO178, PRO195, PRO228, PRO301, PRO302, PRO532, PRO724, PRO730, PRO734, PRO793, PRO871, PRO938, PRO1012, PRO1120, PRO1139, PRO1198, PRO1287, PRO1361, PRO1864, PRO1873, PRO2010, PRO3579, PRO4313, PRO4527, PRO4538, PRO4553, PRO4995, PRO5730, PRO6008, PRO7223, PRO7248 or PRO7261 polypeptide, or agonist thereof, wherein tube formation in said endothelial cell is induced.



1/392

**FIGURE 1**

GCCCACGCGTCCGATGGCGTTCACGTTTCGCGGCCTTCTGCTACATGCTGGCGCTGCTGCT  
CACTGCCGCGCTCATCTTCTTCGCCATTTGGCACATTATAGCATTTGATGAGCTGAAGAC  
TGATTACAAGAATCCTATAGACCAGTGTAATACCCTGAATCCCCCTTGTAAGTCCCAGAGTA  
CCTCATCCACGCTTTCTTCTGTGTCATGTTTCTTTGTGCAGCAGAGTGGCTTACACTGGG  
TCTCAATATGCCCCCTCTTGGCATATCATATTTGGAGGTATATGAGTAGACCAGTGATGAG  
TGGCCCAGGACTCTATGACCCCTACAACCATCATGAATGCAGATATTCTAGCATATTGTCA  
GAAGGAAGGATGGTGCAAATTAGCTTTTTATCTTCTAGCATTTTTTTTACTACCTATATGG  
CATGATCTATGTTTTGGTGAGCTCTTAGAACAACACACAGAAGAATTGGTCCAGTTAAGT  
GCATGCAAAAAGCCACCAAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGC  
CTGTGGAATCTGATCAGTTACTTTAAAAAATGACTCCTTATTTTTTAAATGTTTCCACAT  
TTTTGCTTGTGGAAAGACTGTTTTTCATATGTTTATACTCAGATAAAGATTTTTAAATGGTAT  
TACGTATAAATTAATATAAAATGATTACCTCTGGTGTGACAGGTTTGAACCTTGCACTTC  
TTAAGGAACAGCCATAATCCTCTGAATGATGCATTAATTACTGACTGTCCTAGTACATTG  
GAAGCTTTTGTATTATAGGAACCTGTAGGGCTCATTTTGGTTTCATTGAAACAGTATCTAA  
TTATAAATTAGCTGTAGATATCAGGTGCTTCTGATGAAGTGAAAATGTATATCTGACTAG  
TGGGAAACTTCATGGGTTTCCTCATCTGTGATGTCGATGATTATATATGGATACATTTAC  
AAAAATAAAAAGCGGGAATTTTCCCTTCGCTTGAATATTATCCCTGTATATTGCATGAAT  
GAGAGATTTCCCATATTTCCATCAGAGTAATAAATATACTTGCTTTAATTCTTAAGCATA  
AGTAAACATGATATAAAAATATATGCTGAATTACTTGTGAAGAATGCATTTAAAGCTATT  
TTAAATGTGTTTTTATTTGTAAGACATTACTTATTAAGAAATTGGTTATTATGCTTACTG  
TTCTAATCTGGTGGTAAAGGTATTCTTAAGAAATTGCAGGTACTACAGATTTTCAAACT  
GAATGAGAGAAAATTGTATAACCATCCTGCTGTTCTTTAGTGCAATACAATAAACTCT  
GAAATTAAGACTC

2/392

**FIGURE 2**

MAFTFAAFCYMLALLLTAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNPLVLPEYLIHA  
FFCVMFLCAAEWLTLGLNMPLLAYHIWRYMSRPVMSGPLYDPTTIMNADILAYCQKEGW  
CKLAFYLLAFFYYLYGMIYVLVSS

**Important features:**

**Signal peptide:**

amino acids 1-20

**Type II transmembrane domain:**

amino acids 11-31

**Other transmembrane domain:**

amino acids 57-77 and 123-143

3/392

**FIGURE 3**

GGCTCAGAGGCCCACTGGACCCTCGGCTCTTCCTTGGACTTCTTGTGTGTTCTGTGAGC  
TTCGCTGGATTCAAGGTCTTGGGCATCAGAGGTGAGAGGGTGGGAAGGTCCGCCGCGATG  
GGGAAGCCCTGGCTGCGTGCGCTACAGCTGCTGCTCCTGCTGGGCGCGTCGTGGGCGCGG  
GCGGGCGCCCCGCGCTGCACCTACACCTTCGTGCTGCCCCCGCAGAAGTTCACGGGCGCT  
GTGTGCTGGAGCGGCCCCGCATCCACGCGGGCGACGCCCCGAGGCCGCCAACGCCAGCGAG  
CTGGCGGCGCTGCGCATGCGCGTGGGCCGCCACGAGGAGCTGTTACGCGAGCTGCAGAGG  
CTGGCGGCGGCCGACGGCGCCGTGGCCGGCGAGGTGCGCGCGCTGCGCAAGGAGAGCCGC  
GGCCTGAGCGCGCGCCTGGGCCAGTTGCGCGCGCAGCTGCAGCACGAGGCGGGGCCCGGG  
GCGGGCCCCGGGGCGGATCTGGGGGCGGAGCCTGCCGCGGCGCTGGCGCTGCTCGGGGAG  
CGCGTGCTCAACGCGTCCGCCGAGGCTCAGCGCGCAGCCGCCCGGTTCCACCAGCTGGAC  
GTCAAGTTCGCGAGCTGGCGCAGCTCGTCACCCAGCAGAGCAGTCTCATCGCCCGCCTG  
GAGCGCCTGTGCCCCGGGAGGCGGGCGGGCAGCAGCAGGTCTGCCGCCACCCCCACTG  
GTGCCTGTGGTTCCGGTCCGTCTTGTGGGTAGCACACAGTGACACCAGTAGGATGCTGGAC  
CCAGCCCCAGAGCCCCAGAGAGACCAGACCCAGAGACAGCAGGAGCCCATGGCTTCTCCC  
ATGCCTGCAGGTACCCCTGCGGTCCCCACCAAGCCTGTGGGCCCCGTGGCAGGATTGTGCA  
GAGGCCCGCCAGGCAGGCCATGAACAGAGTGAGTGTATGAACTGCGAGTGGGCCGTAC  
GTAGTGTCAGTATGGTGTGAGCAGCAACTGGAGGGTGGAGGCTGGACTGTGATCCAGCGG  
AGGCAAGATGGTTCAGTCAACTTCTTCACTACCTGGCAGCACTATAAGGCGGGCTTTGGG  
CGGCCAGACGGAGAATACTGGCTGGGCCTTGAACCCGTGTATCAGCTGACCAGCCGTGGG  
GACCATGAGCTGCTGGTTCTCCTGGAGGACTGGGGGGGCGTGGAGCACGTGCCCACTAT  
GATGGCTTCTCCCTGGAACCCGAGAGCGACCACTACCGCCTGCGGCTTGGCCAGTACCAT  
GGTGATGCTGGAGACTCTCTTCTTGGCACAATGACAAGCCCTTCAGCACCGTGGATAGG  
GACCGAGACTCCTATTCTGGTAACTGTGCCCTGTACCAGCGGGGAGGCTGGTGGTACCAT  
GCCTGTGCCCCACTCCAACCTCAACGGTGTGTGGCACCACGGCGGCCACTACCGAAGCCGC  
TACCAGGATGGTGTCTACTGGGCTGAGTTTCGTGGTGGGGCATATTCTCTCAGGAAGGCC  
GCCATGCTCATTCCGGCCCCCTGAAGCTGTGACTCTGTGTTCTCTGTCCCCCTAGGCCCTAG  
AGGACATTGGTCAAGCAGGAGCCCCAAGTTGTTCTGGCCACACCTTCTTTGTGGCTCAGTGC  
CAATGTGTCCACAGAACTTCCCACTGTGGATCTGTGACCCTGGGCGCTGAAAATGGGAC  
CCAGGAATCCCCCGTCAATATCTTGGCCTCAGATGGCTCCCCAAGGTATTTCATATCT  
CGGTTTGAGCTCATATCTTATAATAACACAAAGTAGCCAC

4/392

**FIGURE 4**

MGKPWLRALQLLLLLGASWARAGAPRCTYTFVLPPQKFTGAVCWSGPASTRATPEANAS  
ELAALRMVRGRHEELLRELQRLAAADGAVAGEVRALRKESRGLSARLGQLRAQLQHEAGP  
GAGPGADLGAEPAAALALLGERVLNASAEQAARFHHQLDVKFRELAQLVTQQSSLIAR  
LERLCPPGAGGQQQVLPPPPLVPVVPVRLVGSTSDTSRMLDPAPEPQRDQTORQQEPMAS  
PMPAGHPAVPTKPVGPWQDCAEARQAGHEQSGVYELRVGRHVSVWCEQQLEGGGWTVIQ  
RRQDGSVNFFTTWQHYKAGFGRPDGEYWLGLEPVYQLTSRGDHELLVLLEDWGGRGARAH  
YDGFSLPEPSDHYRLRLGQYHGDAGDSLWHNDKPFSTVDRDRDSYSGNCALYQRGGWY  
HACAHSNLNGVWHHGGHYRSRYQDGVYWAEFRGGAYSLRKAAMLIRPLKL

**Signal peptide:**

Amino acids 1-20

**N-glycosylation sites:**

Amino acids 58-62;145-149

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 97-101

**Tyrosine kinase phosphorylation site:**

Amino acids 441-448

**N-myristoylation sites:**

Amino acids

16-22;23-29;87-93;108-114;121-127;125-131;129-135;187-193;293-299;353-359;378-384;445-451;453-459

**Cell attachment sequence:**

Amino acids 340-343

**Fibrinogen beta and gamma chains C-terminal domain signature:**

Amino acids 418-431

5/392

**FIGURE 5**

CCCACGCGTCCGGCGCCGTGGCCTCGCGTCCATCTTTGCCGTTCTCTCGGACCTGTCACA  
AAGGAGTCGCGCCGCCGCCGCCGCCCTCCCTCCGGTGGGCCCCGGGAGGTAGAGAAAGT  
CAGTGCCACAGCCCGACCGCGCTGCTCTGAGCCCTGGGCACGCGGAACGGGAGGGAGTCT  
GAGGGTTGGGGACGTCTGTGAGGGAGGGGAACAGCCGCTCGAGCCTGGGGCGGGCGGACC  
GGACTGGGGCCGGGGTAGGCTCTGGAAAGGGCCCCGGGAGAGAGGTGGCGTTGGTCAGAAC  
CTGAGAAAACAGCCGAGAGGTTTTCCACCGAGGCCCGCGCTTGAGGGATCTGAAGAGGTTT  
CTAGAAGAGGGTGTTCCCTCTTTCCGGGGTCTCACCAGAAGAGGTTCTTGGGGGTGCGC  
CTTCTGAGGAGGCTGCGGCTAACAGGGCCCAGAACTGCCATTGGATGTCCAGAATCCCCT  
GTAGTTGATAATGTTGGGAATAAGCTCTGCAACTTTCTTTGGCATTTCAGTTGTTAAAAAC  
AAATAGGATGCAAATTCCTCAACTCCAGGTTATGAAAACAGTACTTGGAAAACGAAAAC  
TACCTAAATGATCGTCTTTGGTTGGGCCGTGTTCTTAGCGAGCAGAAGCCTTGGCCAGGG  
TCTGTTGTTGACTCTCGAAGAGCACATAGCCCACTTCCTAGGGACTGGAGGTGCCGCTAC  
TACCATGGGTAATTCCTGTATCTGCCGAGATGACAGTGGAACAGATGACAGTGTGACAC  
CCAACAGCAACAGGCCGAGAACAGTGCAGTACCCACTGCTGACACAAGGAGCCAACCACG  
GGACCCTGTTCCGGCCACCAAGGAGGGGCGGAGGACCTCATGAGCCAAGGAGAAAAGAAACA  
AAATGTGGATGGGCTAGTGTTGGACACACTGGCAGTAATACGGACTCTTGTAAGATAAGTA  
AGTATCTGACTCACGGTCACCTCCAGTGGAATGAAAAGTGTTCTGCCCGGAACCATGACT  
TTAGGACTCCTTCAGTTCTTTAGGACATACTCGCCAAGCCTTGTTGCTCACAGGGCAAAG  
GAGAATATTTAATGCTCCGCTGATGGCAGAGTAAATGATAAGATTTGATGTTTTTGCTT  
GCTGTCATCTACTTTGTCTGGAAATGTCTAAATGTTTCTGTAGCAGAAAACACGATAAAG  
CTATGATCTTTATTAGAG

6/392

## FIGURE 6

MIVFGWAVFLASRSLGQGLLLTLEEHIAHFLGTGGAATTMGNSCICRDDSGTDDSVDTQQ  
QQAENSAVPTADTRSQPRDPVRPPRRGRGPHEPRRKKQNV DGLVLDTLAVIRTLVDKO

Signal peptide:  
amino acids 1-16

Casein kinase II phosphorylation site:  
amino acids 22-26, 50-54, 113-117

N-myristoylation site:  
amino acids 18-24, 32-38, 34-40, 35-41, 51-57

7/392

**FIGURE 7**

CGGACGCGTGGGGGAAACCCCTTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAG  
GGGAACAAGATGGCGGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAACTGGGGCTCCCG  
CCGCTGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTCTGGGGACCGCTTCGGCTGAAGCA  
TTTGA<sup>1</sup>CTCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCC  
TTGCACACCTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTT  
TCAATTTGTCAGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAA  
TCTGCATGTACAGAAGCATATTC<sup>2</sup>CCAATCTGATGAGCAATATGCTTGCCATCTTG<sup>3</sup>TTGC  
CAGAATCAGCTGCCATTCTGCTGAACTGAGACAAGAACA<sup>4</sup>ACTTATGTCCCTGATGCCAAAA  
ATGCACCTACTCTTTCTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCC  
GCACAGAGCTTCATAACCTCTTCATGGACTTTTTATCTTCAAGCCGATGACGGAAAAATA  
GTTATATTCCAGTCTAAGCCAGAAATCCAGTACGCACCACATTTGGAGCAGGAGCCTACA  
AATTTGAGAGAATCATCTCTAAGCAAAATGTCCTATCTGCAAATGAGAAATTCACAAGCG  
CACAGGAATTTTCTTGAAGATGGAGAAAGTGATGGCTTTTTAAGATGCCTCTCTCTTAAC  
TCTGGGTGGATTTTAACTACA<sup>5</sup>ACTCTTGTCCTCTCGGTGATGGTATTGCTTTGGATTTGT  
TGTGCAACTGTTGCTACAGCTGTGGAGCAGTATGTTCCCTCTGAGAAGCTGAGTATCTAT  
GGTGACTTGGAGTTTATGAATGAACAAAAGCTAAACAGATATCCAGCTTCTTCTCTTG<sup>6</sup>TG  
GTTGTTAGATCTAA<sup>7</sup>AACTGAAGATCATGAAGAAGCAGGGCCTCTACCTACAAAAGTGAAT  
CTTGCTCATTCTGAAATTTAAGCATTTTTCTTTTAAAGACAAGTGTAAATAGACATCTAA  
AATTC<sup>8</sup>ACTCCTCATAGAGCTTTTAA<sup>9</sup>AATGGTTTCATTGGATATAGGCCTTAAGAAATCA  
CTATAAAATGCAAATAAAGTTACTCAAATCTGTG

8/392

**FIGURE 8**

MAAPKGSLWVRTQLGLPPLLLLTMALAGGSGTASAEAFDSVLGDTASCHRAQCLTYPLHT  
YPKEEELYACQRCRLFSICQFVDDGIDLNRKLECESACTEAYSQSDEQYACHLGCONQ  
LPFAELRQEQQLMSLMPKMHLFPPLTLVRSFWSMDMSAQSFITSSWTFYLQADDGKIVIF  
QSKPEIQYAPHLEQEPTNLRESSLSKMSYLMRNSQAHRNFLEDGESDGFLRCLSLNSGW  
ILTTTLVLVSVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVR  
SKTEDHEEAGPLPTKVNLAHSEI

**Important features:**

**Signal peptide:**

amino acids 1-31

**Transmembrane domain:**

amino acids 241-260

**N-glycosylation site:**

amino acids 90-93



9/392

**FIGURE 9**

TATTTACCATATCAGATTCACATTCAGTCCTCAGCAAAATGAAGGGCTCCATTTTCACTC  
TGTTTTTATTCTCTGTCTATTTGCCATCTCAGAAGTGCGGAGCAAGGAGTCTGTGAGAC  
TCTGTGGGCTAGAATACATACGGACAGTCATCTATATCTGTGCTAGCTCCAGGTGGAGAA  
GGCATCTGGAGGGGATCCCTCAAGCTCAGCAAGCTGAGACAGGAAACTCCTTCCAGCTCC  
CACATAAACGTGAGTTTTCTGAGGAAAATCCAGCGCAAAACCTTCCGAAGGTGGATGCCT  
CAGGGGAAGACCGTCTTTGGGGTGGACAGATGCCCACTGAAGAGCTTTGGAAGTCAAAGA  
AGCATTCAAGTATGTCAAGACAAGATTTACAACTTTGTGTTGCACTGATGGCTGTTCCA  
TGAATGATTTGAGTGCTCTTTGCTAAGACAAGAGCAAATACCCAATGGGTGGCAGAGCTT  
TATCACATGTTTAATTACAGTGTTTTACTGCCTGGTAGAACACTAATATTGTGTTATTAA  
AATGATGGCTTTTGGGTAGGCAAACTTCTTTTCTAAAAGGTATAGCTGAGCGGTTGAAA  
CCACAGTGATCTCTATTTTCTCCCTTTGCCAAGGTTAATGAACTGTTCTTTCAAATTCT  
ACTAATGCTTTGAAATTTCAAATGCTGCGCAAAATTGCAATAAAAATGCTATAAA

10/392

## FIGURE 10

MKGSIFTLFLFSVLFAISEVRSKESVRLCGLEYIRTVIYICASSRWRRLHLEGIPQAQQAE  
TGNSFQLPHKREFSEENPAQNLPKVDASGEDRLWGGQMPTEELWKSCKHVSMSRQDLQTL  
CCTDGCSMTDLSALC

**Important features:**

**Signal sequence:**

amino acids 1-18

**cAMP- and cGMP-dependent protein kinase phosphorylation  
site:**

amino acids 107-111

**N-myristoylation sites:**

amino acids 3-9, 52-58, 96-102, 125-131

**Insulin family signature:**

amino acids 121-136

**Insulin family proteins:**

amino acids 28-46

11/392

**FIGURE 11**

CCCACGCGTCCGGACAAACTGGAGGTGAAAGGAGCTGGTACTGTCCACTGTGCTGTCCGGT  
GCTGAACCTGAGACGCGAGCGGACCAGTTGCTCCAGCACCTGAAGGCAACGCCCTCTTGC  
ACCCCTCTGTGCCCTGTGGGACCCGCTTCACCAACAGGACCCATATCAACTTGACAAAGGA  
GTGTGGTATCGGACGTGGGAGAGAGTCTCTGTTTGCCACCTGGGCGCTCATTTCAGGCGT  
GACTTTGGAGATTTCTATAGTTTTAGACCAAATATTTTTTTTTTCCCAGCTAAGACGAT  
CTTTTGAGAGTTTTTTTTTTTTTATTGTGATTTATATTTCCACAGCGTTTAGGAATCTTCT  
GGGGGACTTTTGTGACTGTTAAAATAAGGTGAAAAGCAATAAGGATGTTTAAGTGCTGGT  
CAGTTGTCTTGGTTCTCGGATTCATTTTTCTGGAGTCGGAAGGAAGGCCAACCAAAGAAG  
GAGGATATGGCCTTAAATCCTATCAGCCTCTAATGAGATTGCGACATAAGCAGGAAAAAA  
ATCAAGAAAGTTCAAGAGTCAAAGGATTTATGATTTCAGGATGGCCCTTTTGGATCTTGTG  
AAAATAAGTACTGTGGTTTGGGAAGACACTGTGTTACCAGCAGAGAGACAGGGCAAGCAG  
AATGTGCCTGTATGGACCTTTGCAAACGTCACCTACAAACCTGTGTGTGGATCTGACGGAG  
AATTCTATGAAAACCACTGTGAAGTGCACAGAGCTGCTTGCCCTGAAAAAACAAAAGATTA  
CCATTGTTACAAATGAAGACTGCTTCTTTAAAGGAGATAAGTGCAAGACTACTGAATACA  
GCAAGATGAAAAATATGCTATTAGATTTACAAAATCAAAAATATATTATGCAAGAAAATG  
AAAATCCTAATGGCGACGACATATCTCGGAAGAAGCTATTGGTGGATCAAATGTTTAAAT  
ATTTTGATGCAGACAGTAATGGACTTGTAGATATTAATGAACTAACTCAGGTGATAAAAC  
AGGAAGAACTTGGCAAGGATCTCTTTGATTGTACTTTGTATGTTCTATTGAAATATGATG  
ATTTTAATGCTGACAAGCACCTGGCTCTTGAAGAATTTTATAGAGCATTCCAAGTGATCC  
AGTTGAGTCTGCCAGAAGATCAGAACTAAGCATCACTGCAGCAACTGTGGGACAAAGTG  
CTGTTCTGAGCTGTGCCATTCAAGGAACCTGAGACCTCCCATTATCTGGAAAAGGAACA  
ATATTATTCTAAATAATTTAGATTTGGAAGACATCAATGACTTTGGAGATGATGGGTCCT  
TGTATATTACTAAGGTTACCACAACCTCACGTTGGCAATTACACCTGCTATGCAGATGGCT  
ATGAACAAGTCTATCAGACTCACATCTTCCAAGTGAATGTTCTCCAGTCATCC

12/392

**FIGURE 12**

MFKCWSVVLVLGFIFLESEGRPTKEGGYGLKSYQPLMRLRHKQEKQESSRVKGFMIQDG  
PFGSCENKYCGLGRHCVTSRETGQAECACMDLCKRHYKPVCSDGEFYENHCEVHRAACL  
KKQKITIVHNEDCFFKGDCKTTEYSKMKNMLLDLQNQKYIMQENENPNGDDISRKKLLV  
DQMFKYFDADSNGLVDINELTQVIKQEELGKDLFDCTLYVLLKYDDFNADKHLALEEFYR  
AFQVIQLSLPEDQKLSITAATVGQSAVLSCAIQGTLRPPIIWKRNNIILNNDLEDINDF  
GDDGSLYITKVTTHVGNVTCYADGYEQVYQTHIFQVNVPPVI

**Signal sequence:**

Amino acids 1-20

**N-glycosylation site:**

Amino acids 318-322

**Tyrosine kinase phosphorylation sites:**

Amino acids 21-29;211-220

**N-myristoylation sites:**

Amino acids 63-69;83-89;317-323

**Prokaryotic membrane lipoprotein lipid attachment site:**

Amino acids 260-271

13/392

**FIGURE 13**

TGCCGGGCTGCGGGGCGCCTTGACTCTCCCTCCACCCTGCCTCCTCGGGCTCCACTCGTC  
TGCCCCCTGGACTCCCGTCTCCTCCTGTCTCCGGCTTCCCAGAGCTCCCTCCTTATGGCA  
GCAGCTTCCCGCGTCTCCGGCGCAGCTTCTCAGCGGACGACCCTCTCGCTCCGGGGCTGA  
GCCCAGTCCCTGGATGTTGCTGAAACTCTCGAGATCATGCGCGGGTTTGGCTGCTGCTTC  
CCGCGCGGGTGCCACTGCCACCGCCCGCCGCTCTGCTGCCGCGGTCCGCGGGATGCTCAG  
TAGCCCGCTGCCCGGCCCCCGCGATCCTGTGTTCTCGGAAGCCGTTTGCTGCTGCAGAG  
TTGCACGAAC TAGTCATGGTGCTGTGGGAGTCCCCGCGGCAGTGCAGCAGCTGGACACTT  
TGCGAGGGCTTTTGCTGGCTGCTGCTGCTGCCCGTCATGCTACTCATCGTAGCCCCGCCG  
GTGAAGCTCGCTGCTTTCCCTACCTCCTTAAGTGACTGCCAAACGCCACCGGCTGGAAT  
TGCTCTGGTTATGATGACAGAGAAAATGATCTCTTCCCTCTGTGACACCAACACCTGTAAA  
TTTGATGGGGAATGTTTAAGAATTGGAGACACTGTGACTTGCGTCTGTCAAGTCAAGTGC  
AACATGACTATGTGCCTGTGTGTGGCTCCAATGGGGAGAGCTACCAGAATGAGTGTAC  
CTGCGACAGGCTGCATGCAAACAGCAGAGTGAGATACTTGTGGTGTGAGAAGGATCATGT  
GCCACAGATGCAGGATCAGGATCTGGAGATGGAGTCCATGAAGGCTCTGGAGAACTAGT  
CAAAAGGAGACATCCACCTGTGATATTTGCCAGTTTGGTGCAGAATGTGACGAAGATGCC  
GAGGATGTCTGGTGTGTGTGTAATATTGACTGTTCTCAAACCAACTTCAATCCCCTCTGC  
GCTTCTGATGGGAAATCTTATGATAATGCATGCCAAATCAAAGAAGCATCGTGTGAGAAA  
CAGGAGAAAATTGAAGTCATGTCTTTGGGTCGATGTCAAGATAACACAACACTACAACACT  
AAGTCTGAAGATGGGCATTATGCAAGAACAGATTATGCAGAGAATGCTAACAAATTAGAA  
GAAAGTGCCAGAGAACACCACATACCTTGTCCGGAACATTACAATGGCTTCTGCATGCAT  
GGGAAGTGTGAGCATTCTATCAATATGCAGGAGCCATCTTGCAGGTGTGATGCTGGTTAT  
ACTGGACAACACTGTGAAAAAAGGACTACAGTGTTCTATACGTTGTTCCCGGTCCTGTA  
CGATTTTCAGTATGTCTTAATCGCAGCTGTGATTGGAACAATTCAGATTGCTGTCTGT  
GTGGTGGTCCTCTGCATCACAAGGAAATGCCCCAGAAGCAACAGAATTCACAGACAGAAG  
CAAAATACAGGGCACTACAGTTCAGACAATACAACAAGAGCGTCCACGAGGTTAATCTAA  
AGGGAGCATGTTTCACAGTGGCTGGACTACCGAGAGCTTGGACTACACAATACAGTATTA  
TAGACAAAAGAATAAGACAAGAGATCTACACATGTTGCCTTGCATTTGTGGTAATCTACA  
CCAATGAAAACATGTACTACAGCTATATTTGATTATGTATGGATATATTTGAAATAGTAT  
ACATTGTCTTGATGTTTTTTCTGTAATGTAAATAAACTATTTATATCACACAATATAGTT  
TTTTCTTCCCATGTATTTGTTATATATAATAAATACTCAGTGATGAG

14/392

**FIGURE 14**

MVLWESPRQCSSWTLCGFCWLLLLPVMLLIVARPVKLAAFPSTLSDCQTPTGWNC SGY  
DDRENDLFLCDTNTCKFDGECLRIGDTVTCVCQFKCNNDYVPVCGSNGESYQNECYLRQ  
AACKQQSEILVVSEGSCATDAGSGSGDGVHEGSGETSQKETSTCDICQFGAECDEDAED  
VWCVCNIDCSQTNFNPLCASDGKSYDNACQIKEASCQKQEKIEVMSLGRCQDNTTTT TK  
SEDGHYARTDYAENANKLEESAREHHIPCPEHYNGFCMHGKCEHSINMQEPSCRC DAGY  
TGQHCEKKDYSVLYVVP GPVRFQYVLIAAVIGTIQIAVICVVVLCITRKCPRSNRIHRQ  
KQNTGHYSSDNTTRASTRLI

15/392

**FIGURE 15**

GGAGCTCAGCCCAGTGGGCAGTCTGAAGATGGCCAATTACACGCTGGCACCAGAGGATGA  
ATATGATGTCCTCATAGAAGGTGAACTGGAGAGCGATGAGGCAGAGCAATGTGACAAGTA  
TGACGCCCAGGCACTCTCAGCCCAGCTGGTGCCATCACTCTGCTCTGCTGTGTTTGTGAT  
CGGTGTCCTGGACAATCTCCTGGTTGTGCTTATCCTGGTAAAATATAAAGGACTCAAACG  
CGTGAAAAATATCTATCTTCTAAACTTGGCAGTTTCTAACTTGTGTTTCTTGCTTACCCT  
GCCCTTCTGGGCTCATGCTGGGGGCGATCCCATGTGTAAAATTCTCATTGGACTGTACTT  
CGTGGGCCTGTACAGTGAGACATTTTTCAATTGCCTTCTGACTGTGCAAAGGTACCTAGT  
GTTTTTGCACAAGGGCAACTTTTTCTCAGCCAGGAGGAGGGTGCCCTGTGGCATCATTAC  
AAGTGTCTTGGCATGGGTAAACAGCCATTCTGGCCACTTTGCCTGAATACGTGGTTTATAA  
ACCTCAGATGGAAGACCAGAAATACAAGTGTGCATTTAGCAGAACTCCCTTCCTGCCAGC  
TGATGAGACATTCTGGAAGCATTTTTCTGACTTTAAAAATGAACATTTCCGTTCTTGTCCT  
CCCCCTATTTATTTTTTACATTTCTCTATGTGCAAATGAGAAAAACACTAAGGTTTCAGGA  
GCAGAGGTATAGCCTTTTCAAGCTTGTTTTTGCCATAATGGTAGTCTTCCCTTCTGATGTG  
GGCGCCCTACAATATTGCATTTTTCTGTCCACTTTCAAAGAACACTTCTCCCTGAGTGA  
CTGCAAGAGCAGCTACAATCTGGACAAAAGTGTTTACATCACTAAACTCATCGCCACCAC  
CCACTGCTGCATCAACCCTCTCCTGTATGCGTTTCTTGATGGGACATTTAGCAAATACCT  
CTGCCGCTGTTTCCATCTGCGTAGTAACACCCCACTTCAACCCAGGGGGCAGTCTGCACA  
AGGCACATCGAGGGAAGAACCTGACCATTCCACCGAAGTGTA~~AA~~ACTAGCATCCACCAAAT  
GCAAGAAGAATAAACATGGATTTTCATCTTCTGCATTATTTTCATGTAAATTTTCTACAC  
ATTTGTATACAAAATCGGATACAGGAAGAAAAGGGAGAGGTGAGCTAACATTTGCTAAGC  
ACTGAATTTGTCTCAGGCACCGTGCAAGGCTTTTACAAACGTGAGCTCCTTCGCCTCCT  
ACCACTTGTCCATAGTGTGGATAGGACTAGTCTCATTTCTCTGAGAAGAAAATAAGGCG  
CGGAAATTTGTCTAAGATCACATAACTAGGAAGTGGCAGAACTGATTCTCCAGCCCTGGT  
AGCATTTGCTCAGAGCCTACGCTTGGTCCAGAACATCAAACCTCAAACCCTGGGGACAAA  
CGACATGAAATAAATGTATTTTAAAACATCTAAAA

16/392

**FIGURE 16**

MANYTLAPEDEYDVLIEGELESDEAEQCDKYDAQALSAQLVPSLCSAVFVIGVLDNLLVV  
LILVKYKGLKRVENIYLLNLAVSNLCFLLTLPFWAHAGGDPCKILIGLYFVGLYSETFF  
NCLLTVQRYLVFLHKGNNFSARRRVPCGIITSVLAWVTAILATLPEYVVYKPMEDQKYK  
CAFSRTPFLPADETFWKHFLTLKMNISVLVLPLFIFTFLYVQMRKTLRFREQRYSLFKLV  
FAIMVVFLLMWAPYNIAFFLSTFKEHFSLSDCKSSYNLDKSVHITKLIATTHCCINPLLY  
AFLDGTFSKYLRCFHLRSNTPLQPRGQSAQGTSREEPDHSTEV

**Signal sequence:**

None

**Transmembrane domain:**

41-61, 76-96, 109-129, 147-167, 199-219, 237-257, 285-305

**7 transmembrane receptor (rhodopsin family):**

55-300

**N-glycosylation site:**

3-6, 205-208

**Tyrosine kinase phosphorylation site:**

70-76, 171-179, 228-234

**N-myristoylation site:**

52-57, 136-141, 148-153

**G-protein coupled receptors:**

55-85, 96-136, 209-220, 235-254, 292-308



17/392

**FIGURE 17**

CGGACGCGTGGGCGGACGCGTGGGCGGCCACGGCGCCCGCGGGCTGGGGCGGTGCGTTCTTCCTTCTCCGTGGCCTACGAGGGTCCCCAGCCTGGGTAAAGATGGCCCCATGGCCCCCGAAGGGCCTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTTCCTCAACCTCCCAGGACCTATCTGGCTCCAGCCCTCTCCACCTCCCCAGTCTTCTCCCCCGCCTCAGCCCCATCCGTGTCATACCTGCCGGGGACTGGTTGACAGCTTTAAACAAGGGCCTGGAGAGAACCATCCGGGGACAACCTTGGAGGTGGAAACACTGCCTGGGAGGAAGAGAATTGTCCAAATACAAAGACAGTGAGACCCGCCTGGTAGAGGTGCTGGAGGGTGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCTGAGTGAGGAGCTGGTGGAGAGCTGGTGGTTTCACAAGCAGCAGGAGGCCCCGGACCTCTTCCAGTGGCTGTGCTCAGATTCCCTGAAGCTCTGCTGCCCCGCAGGCACCTTCGGGCCCTCCTGCCCTTCCCTGTCTGGGGGAACAGAGAGGCCCTGCGGTGGCTAGGGCAGTGTGAAGGAGAAGGGACACGAGGGGCAGCGGGCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCTGTGGCCAGTGTGGCCCTTGGCTACTTTGAGGCAGAACGCAACGCCAGCCATCTGGTATGTTTCGGCTTGTTTGGCCCCCTGTGCCGATGCTCAGGACCTGAGGAATCAAACTGTTTGCAATGCAAGAAGGGCTGGGCCCTGCATCACCTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAACCTGTGGAGCTGACCAATTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAGGCCTGCCTAGGCTGCATGGGGGCAGGGCCAAGTTCGCTGTAAGAAGTGTAGCCCTGGCTATCAGCAGGTGGGCTCCAAGTGTCTCGATGTGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGAAACAAGCAGTGTGAAAACACCGAGGGCGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAGATGGAAGGCATCTGTGTGAAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGACAGAAGACGAGTTGGTGGTGTGTCAGCAGATGTTCTTTGGCATCATCATCTGTGCACTGGCCACGCTGGCTGCTAAGGGCGACTTGGTGTTCACCGCCATCTTCATTGGGGCTGTGGCGGCCATGACTGGCTACTGGTTGTCAAGCGCAGTGACCGTGTGCTGGAGGGCTTCATCAAGGGCAGATAATCGCGGCCACCACCTGTAGGACCTCCTCCCACCCACGCTGCCCCCAGAGCTTGGGCTGCCCTCCTGCTGGACACTCAGGACAGCTTGGTTTATTTTTGAGAGTGGGGTAAGCACCCCTACCTGCCTTACAGAGCAGCCCAGGTACCCAGGCCCCGGGCAGACAAGGCCCTGGGGTAAAAAGTAGCCCTGAAGGTGGATACCATGAGCTCTTCACCTGGCGGGGACTGGCAGGCTTCACAATGTGTGAATTTCAAAAGTTTTTCCTTAATGGTGGCTGCTAGAGCTTTGGCCCCCTGCTTAGGATTAGGTGGTCTCACAGGGGTGGGGCCATCACAGCTCCCTCCTGCCAGCTGCATGCTGCCAGTTCCTGTCTGTGTTTCAACCATCCCCACACCCATTGCCACTTATTTATTCATCTCAGGAAATAAAGAAAGGTCTTGAAAGTTAAAAA

18/392

**FIGURE 18**

MAPWPPKGLVPAVLWGLSLFLNLP GPIWLQPSPPPQSSPPPQPHPCHTCRGLVDSFNKGL  
ERTIRDNFGGGNTAWEEENLSKYKDSETRLVEVLEGVCSKSDFECHRLELSEELVESWW  
FHKQQEAPDLFQWLCSDSLKLCCPAGTFGPSCLPCPGGTERPCGGYGQCEGEGTRGGSGH  
CDCQAGYGGEACGQCGLGYFEAERNASHLVCSACFGPCARCSGP EESNCLQCKKGWALHH  
LKCVDIDECEGTEGANCGADQFCVNTEGSYECRCAKACLGCMGAGPGRCKKCSPGYQQVG  
SKCLDVDECETEVC PGENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTE  
DELVLQQMFFGIIICALATLAAKGDLVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR

19/392

**FIGURE 19**

GCCCCGGGACTGGCGCAAGGTGCCCAAGCAAGGAAAGAAATAATGAAGAGACACATGTGTT  
AGCTGCAGCCTTTTGAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCT  
TTACCACGCTTGTTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCCAGCAT  
GAATCTGGTAGACCTGTGGTTAACCCGTTCCCTCTCCATGTGTCTCCTCCTACAAAGTTT  
TGTTCTTATGATACTGTGCTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTC  
TTCCTCTGGGGGTTTAAATGTACCTGTAGCAATGCAAAATCTCAAGGAAATACCTAGAGA  
TCTTCCTCCTGAAACAGTCTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAA  
TGAAATTTTTTAAGGACCTCCATCAACTGAGAGTTCTCAACCTGTCCAAAAATGGCATTGA  
GTTTATCGATGAGCATGCCTTCAAAGGAGTAGCTGAAACCTTGCAGACTCTGGACTTGTC  
CGACAATCGGATTCAAAGTGTGCACAAAAATGCCTTCAATAACCTGAAGGCCAGGGCCAG  
AATTGCCAACAACCCCTGGCACTGCGACTGTACTCTACAGCAAGTTCTGAGGAGCATGGC  
GTCCAATCATGAGACAGCCCACAACGTGATCTGTAAAACGTCCGTGTTGGATGAACATGC  
TGGCAGACCATTCTCAATGCTGCCAACGACGCTGACCTTTGTAACCTCCCTAAAAAAC  
TACCGATTATGCCATGCTGGTCACCATGTTTGGCTGGTTCACTATGGTGATCTCATATGT  
GGTATATTATGTGAGGCAAAATCAGGAGGATGCCCGGAGACACCTCGAATACTTGAAATC  
CCTGCCAAGCAGGCAGAAGAAAGCAGATGAACCTGATGATATTAGCACTGTGGTATAGTG  
TCCAAACTGACTGTCAATTGAGAAAGAAAGAAAGTAGTTTGCGATTGCAGTAGAAATAAGT  
GGTTTACTTCTCCCATCCATTGTAAACATTTGAAACTTTGTATTTTCAAGTTTTTTTGAAT  
TATGCCACTGCTGAACTTTTAAACAAACACTACAACATAAATAATTTGAGTTTAGGTGATC  
CACCCCTTAATTGTACCCCGATGGTATATTTCTGAGTAAGCTACTATCTGAACATTAGT  
TAGATCCATCTCACTATTTAATAATGAAATTTATTTTTTTAATTTAAAAGCAAATAAAAG  
CTTAACTTTGAACCATGGGAAAAAAAAAAAAAAAAAAAAAAAAACA

20/392

**FIGURE 20**

MNLVDLWLTRSLSMCLLLQSFVLMILCFHSASMC PKGCLCSSSGGLNVTCSNANLKEIPR  
DLPPETVLLYLD SNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGVAETLQTLDL  
SDNRIQSVHKNAFNNLKARARIANNPWHCDCTLQQVLRSMASNHETAHN VICKTSVLDEH  
AGRPFLNAANDADLCNLPKKT TDYAMLVTMFGWFTMVISYV VYYVRQNQEDARRHLEYLK  
SLPSRQKKADEPDDISTVV

**Signal sequence:**  
amino acids 1-33

**Transmembrane domain:**  
amino acids 205-220

**N-glycosylation site:**  
amino acids 47-51, 94-98

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**  
amino acids 199-203

**Casein kinase II phosphorylation site:**  
amino acids 162-166, 175-179

**N-myristoylation site:**  
amino acids 37-43, 45-51, 110-116

21/392

**FIGURE 21**

CGCCACCACTGCGGCCACCGCCAATGAAACGCCTCCCGCTCCTAGTGGTTTTTTCCACTT  
TGTTGAATTGTTTCTTATACTCAAAATTGCACCAAGACACCTTGTCTCCCAAATGCAAAAT  
GTGAAATACGCAATGGAATTGAAGCCTGCTATTGCAACATGGGATTTTCAGGAAATGGTG  
TCACAATTTGTGAAGATGATAATGAATGTGGAAATTTAACTCAGTCCTGTGGCGAAAATG  
CTAATTGCACTAACACAGAAGGAAGTTATTATTGTATGTGTGTACCTGGCTTCAGATCCA  
GCAGTAACCAAGACAGGTTTATCACTAATGATGGAACCGTCTGTATAGAAAATGTGAATG  
CAAACCTGCCATTTAGATAATGTCTGTATAGCTGCAAAATATTAATAAACTTTAACAAAA  
TCAGATCCATAAAAGAACCTGTGGCTTTGCTACAAGAAGTCTATAGAAATTCTGTGACAG  
ATCTTTTACCAACAGATATAATTACATATATAGAAATATTAGCTGAATCATCTTCATTAC  
TAGGTTACAGAACAACACTATCTCAGCCAAGGACACCCTTTCTAACTCAACTCTTACTG  
AATTTGTAAAAACCGTGAATAAATTTGTTCAAAGGGATACATTGTAGTTTGGGCAAGT  
TATCTGTGAATCATAGGAGAACACATCTTACAAAACCTCATGCACACTGTTGAACAAGCTA  
CTTTAAGGATATCCCAGAGCTTCCAAAAGACCACAGAGTTTGATACAAATTCACCGGATA  
TAGCTCTCAAAGTTTTCTTTTTTGATTTCATATAACATGAAACATATTCATCCTCATATGA  
ATATGGATGGAGACTACATAAATATATTTCCAAAGAGAAAAGCTGCATATGATTCAAATG  
GCAATGTTGCAGTTGCATTTTTATATTATAAGAGTATTGGTCCTTTGCTTTCATCATCTG  
ACAACCTTCTTATTGAAACCTCAAATTTATGATAATTCTGAAGAGGAGGAAAGAGTCATAT  
CTTCAGTAATTTTCACTCTCAATGAGCTCAAACCCACCCACATTATATGAACTTGAAAAAA  
TAACATTTACATTAAAGTCATCGAAAGGTCACAGATAGGTATAGGAGTCTATGTGCATTTT  
GGAATTACTCACCTGATACCATGAATGGCAGCTGGTCTTCAGAGGGCTGTGAGCTGACAT  
ACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCTGACACATTTTGAATTTTGA  
TGTCCTCTGTCCTTCCATTGGTATTAAAGATTATAATATTCTTACAAGGATCACTCAAC  
TAGGAATAATTATTTCACTGATTTGTCTTGCCATATGCATTTTTTACCTTCTGGTTCTTCA  
GTGAAATTCAAAGCACCAGGACAACAATTCACAAAAATCTTTGCTGTAGCCTATTTCTTG  
CTGAACTTGTTTTTCTTGTGGGATCAATACAAATACTAATAAGCTCTTCTGTTCAATCA  
TTGCCGGAAGTGTACACTACTTCTTTTTAGCTGCTTTTGCATGGATGTGCATTGAAGGCA  
TACATCTCTATCTCATTGTTGTGGGTGTCTATACAACAAGGGATTTTTGCACAAGAATT  
TTTATATCTTTGGCTATCTAAGCCCAGCCGTGGTAGTTGGATTTTCGGCAGCACTAGGAT  
ACAGATATTATGGCACAACCAAGTATGTTGGCTTAGCACCGAAAAACAATTTATTTGGA  
GTTTTATAGGACCAGCATGCCTAATCATTCTTGTTAATCTCTTGGCTTTTGGAGTCATCA  
TATACAAAGTTTTTCGTCACTGCAGGGTTGAAACCAGAAAGTTAGTTGCTTTGAGAAC  
TAAGGTCTTGTGCAAGAGGAGCCCTCGCTCTTCTGTTCTTCTCGGCACCACCTGGATCT  
TTGGGGTTCTCCATGTTGTGCACGCATCAGTGGTTACAGCTTACCTCTTCACAGTCAGCA  
ATGCTTTCCAGGGGATGTTCAATTTTTTATTCCTGTGTGTTTTATCTAGAAAGATTCAAG  
AAGAATATTACAGATTGTTCAAAAATGTCCCTGTTGTTTTGGATGTTTAAGGTAAACAT  
AGAGAATGGTGGATAATTACAACCTGCACAAAAATAAAAAATCCAAGCTGTGGATGACCAA  
TGTATAAAAAATGACTCATCAAATTATCCAATTATTAACCTACTAGACAAAAAGTATTTTAA  
ATCAGTTTTTCTGTTTATGCTATAGGAACTGTAGATAATAAGGTAAAAATTATGTATCATA  
TAGATATACTATGTTTTTCTATGTGAAATAGTTCTGTCAAAAATAGTATTGCAGATATTT  
GGAAAGTAATTGGTTTTCTCAGGAGTGATATCACTGCACCCAAGGAAAGATTTTCTTTCTA  
ACACGAGAAGTATATGAATGTCTGAAGGAAACCACTGGCTTGATATTTCTGTGACTCGT  
GTTGCCTTTGAACTAGTCCCCTACCACCTCGGTAATGAGCTCCATTACAGAAAGTGGA  
CATAAGAGAATGAAGGGCAGAAATATCAAACAGTGAAAAGGGAATGATAAGATGTATTTT  
GAATGAACTGTTTTTCTGTAGACTAGCTGAGAAATTGTTGACATAAAATAAAGAATTGA  
AGAAACACATTTTACCATTTTGTGAATTGTTCTGAACTTAAATGTCCACTAAAACAACCTT  
AGACTTCTGTTTGTCTAAATCTGTTTCTTTTTCTAATATTCTAAAAAAGGTTT  
ACCTCCACAAATTGAAA

22/392

**FIGURE 22**

MKRLPLLVVVFSTLLNCSYTQNCTKTPCLPNAKCEIRNGIEACYCNMGFSGNGVTICEDDN  
ECGNLTQSCGENANCTNTEGSYYCMCVPGFRSSSNQDRFITNDGTVCIENTVNANCHLDNV  
CIAANINKTLTKIRSIKEPVALLQEVYRNSVTDLSPTDIITYIEILAESSSLGYKNNTI  
SAKDTLSNSTLTEFVKTVNNFVQRDTFVVWDKLSVNHRRTHLTCLMHTVEQATLRISQSF  
QKTTEFDTNSTDIALKVFFFDSDYNMKHIHPHMNMDGDYINIFPKRKAAYDSNGNVAVAF  
YYKSIGPLLSSSDNFLLKPQNYDNSEEEERVISSVISVSMSSNPPTLYELEKITFTLSHR  
KVTDRYRSLCAFWNYSPTMNGSWSSEGCETYSNETHTSCRCNHLTHFAILMSSGPSIG  
IKDYNILTRITQLGIIISLICLAICIFTFWFFSEIQSTRTTIHKNLCCSLFLAELVFLVG  
INTNTNKLFCSSIAGLLHYFFLAFAWMCIIEGILYLVVGVYINKGFLHKNFYIFGYLS  
PAVVVGFSAAALGYRYYGTTKVCWLSTENNFIWSFIGPACLIILVNLLAFGVIIYKVRHT  
AGLKPEVSCFENIRSCARGALALLFLLGTTWIFGVLHVHASVVTAYLFTVSNAFQGMFI  
FLFLCVLSRKIQEYYRLFKNPCCFGCLR

23/392

**FIGURE 23**

CTCCTCTTAACATACTTGCAGCTAAACTAAATATTGCTGCTTGGGGACCTCCTTCTAGC  
CTTAAATTTTCAGCTCATCACCTTCACCTGCCTTGGTCATGGCTCTGCTATTCTCCTTGAT  
CCTTGCCATTTGCACCAGACCTGGATTCCCTAGCGTCTCCATCTGGAGTGC GGCTGGTGGG  
GGGCCTCCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGT  
GTGTGATGACGGCTGGGACATTAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGG  
AGCTGCCAGCGGAACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAA  
GGTCCTCATCCAATCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCA  
AGAAGAAGTTTATGATTGTTTACATGATGAAGATGCTGGGGCATCGTGTGAGAAGCCAGA  
GAGCTCTTTCTCCCCAGTCCCAGAGGGTGTGAGGCTGGCTGACGGCCCTGGGCATTGCAA  
GGGACGCGTGGAAGTGAAGCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAG  
CCTCCGGGCGCGAAAGGTGGTGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCA  
AAAACGCTGCAACAAGCATGCCCTATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATG  
CTCAGGACGAGAAGCAACCCTTCAGGATTGCCCTTCTGGGCCTTGGGGGAAGAACACCTG  
CAACCATGATGAAGACACGTGGGTGGAATGTGAAGATCCCTTTGACTTGAGACTAGTAGG  
AGGAGACAACCTCTGCTCTGGGCGACTGGAGGTGCTGCACAAGGGCGTATGGGGCTCTGT  
CTGTGATGACAACCTGGGGAGAAAAGGAGGACCAGGTGGTATGCAAGCAACTGGGCTGTGG  
GAAGTCCCTCTCTCCCTCCTTCAGAGACCGGAAATGCTATGGCCCTGGGGTTGGCCGCAT  
CTGGCTGGATAATGTTTCGTTGCTCAGGGGAGGAGCAGTCCCTGGAGCAGTGCCAGCACAG  
ATTTTGGGGGTTTTACGACTGCACCCACCAGGAAGATGTGGCTGTCATCTGCTCAGTGTA  
GGTGGGCATCATCTAATCTGTTGAGTGCCTGAATAGAAGAAAAACACAGAAGAAGGGAGC  
ATTTACTGTCTACATGACTGCATGGGATGAACACTGATCTTCTTCTGCCCTTGGACTGGG  
ACTTATACTTGGTGCCCTGATTCTCAGGCCTTCAGAGTTGGATCAGAACTTACAACATC  
AGGTCTAGTTCTCAGGCCATCAGACATAGTTTGGAACTACATCACCACCTTTCCTATGTC  
TCCACATTGCACACAGCAGATTCCAGCCTCCATAATTGTGTGTATCAACTACTTAAATA  
CATTTCTCACACACACACACACACACACACACACACACACACATACACCATTGTCC  
TGTTTCTCTGAAGAACTCTGACAAAATACAGATTTTGGTACTGAAAGAGATTCTAGAGGA  
ACGGAATTTTAAGGATAAAATTTTCTGAATTGGTTATGGGGTTTCTGAAATTGGCTCTATA  
ATCTAATTAGATATAAAATTTCTGGTAACTTTATTTACAATAATAAAGATAGCACTATGTG  
TTCAAA

24/392

**FIGURE 24**

MALLFSLILAICTRPGFLASPSGVRLVGGLHRCEGRVEVEQKGQWGTVCDDGWDIKDVAV  
LCRELGCGAASGTPSGILYEPPAEKEQKVLIQSVSCTGTEDTLAQCEQEVEYDCSHDEDA  
GASCENPESSFSPVPEGVRLADGPGHCKGRVEVKHQNQWYTVCTGWSLRAAKVVCRLG  
CGRAVLTKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPWGKNTCNHDEDTWVECED  
PFDLRLVGGDNLCSGRLEVLHKGWGSVCCDNWGEKEDQVCKQLGCGKSLSPSFRDRKC  
YGPVGRIWLDNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICSV

**Signal sequence:**  
amino acids 1-15

**Casein kinase II phosphorylation site:**  
amino acids 47-51, 97-101, 115-119, 209-213, 214-218, 234-238,  
267-271, 294-298, 316-320, 336-340

**N-myristoylation site:**  
amino acids 29-35, 43-49, 66-72, 68-74, 72-78, 98-104, 137-143,  
180-186, 263-269, 286-292

**Amidation site:**  
amino acids 196-200

**Speract receptor repeated domain signature:**  
amino acids 29-67, 249-287



25/392

## FIGURE 25

CGGACGCGTGGGCGTCCGGCGGTTCGACAGCCAGGAGGCGGAGGCGCGCGGGCCAGCCTG  
GGCCCCAGCCACACCTTCACCAGGGCCAGGAGCCACCATGTGGCGATGTCCACTGGGG  
CTACTGCTGTTGCTGCCGCTGGCTGGCCACTTGGCTCTGGGTGCCCAGCAGGGTCGTGGG  
CGCCGGGAGCTAGCACCGGGTCTGCACCTGCGGGGCATCCGGGACGCGGGAGGCCGGTAC  
TGCCAGGAGCAGGACCTGTGCTGCCGCGGCCGTGCCGACGACTGTGCCCTGCCCTACCTG  
GGCGCCATCTGTTACTGTGACCTCTTCTGCAACCGCACGGTCTCCGACTGCTGCCCTGAC  
TTCTGGGACTTCTGCCTCGGCGTGCCACCCCTTTTCCCCCGATCCAAGGATGTATGCAT  
GGAGGTCGTATCTATCCAGTCTTGGAACGTACTGGGACAACGTGAACCGTTGCACCTGC  
CAGGAGAACAGGCAGTGGCATGGTGGATCCAGACATGATCAAAGCCATCAACCAGGGCAA  
TATGGCTGGCAGGCTGGGAACCACAGCGCCTTCTGGGGCATGACCCTGGATGAGGGCAT  
TCGCTACCGCCTGGGCACCATCCGCCCATCTTCTCGGTCTGAACATGCATGAAATTTA  
TACAGTGCTGAACCCAGGGGAGGTGCTTCCACAGCCTTCGAGGCCTCTGAGAAGTGGCC  
CAACCTGATTCATGAGCCTCTTGACCAAGGCAACTGTGCAGGCTCCTGGGCCTTCTCCAC  
AGCAGCTGTGGCATCCGATCGTGTCTCAATCCATTCTCTGGGACACATGACGCCTGTCTCT  
GTGCCCCAGAACCTGCTGTCTTGTGACACCCACCAGCAGCAGGGCTGCCGCGGTGGGCG  
TCTCGATGGTGCCTGGTGGTTCTGCGTCCGCGAGGGGTGGTGTCTGACCATGCTACCC  
CTTCTCGGGCCGTGAACGAGACGAGGCTGGCCCTGCGCCCCCTGTATGATGCACAGCCG  
AGCCATGGGTCCGGGCAAGCGCCAGGCCACTGCCCACTGCCCAACAGCTATGTTAATAA  
CAATGACATCTACCAGGTCACTCCTGTCTACCGCCTCGGCTCCAACGACAAGGAGATCAT  
GAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCATGGAGGTGCATGAGGACTTCTT  
CCTATACAAGGGAGGCATCTACAGCCACAGCCAGTGAGCCTTGGGAGGCCAGAGAGATA  
CCGCCGGCATGGGACCCACTCAGTCAAGATCACAGGATGGGGAGAGGAGACGCTGCCAGA  
TGGAAGGACGCTCAAATACTGGACTGCGGCCAACTCCTGGGGCCAGCCTGGGGCGAGAG  
GGGCCACTTCCGCATCGTGC GCGCGTCAATGAGTGCGACATCGAGAGCTTCGTGCTGGG  
CGTCTGGGGCCGCGTGGGCATGGAGGACATGGGTCACTGAGGCTGCGGGCACACGC  
GGGGTCCGGCCTGGGATCCAGGCTAAGGGCCGGCGGAAGAGGCCCAATGGGGCGGTGAC  
CCCAGCCTCGCCGACAGAGCCCGGGGCGCAGGCGGGCGCCAGGGCGCTAATCCCGGCGC  
GGGTTCCGCTGACGCGAGCGCCCGCCTGGGAGCCGCGGGCAGGCGAGACTGGCGGAGCCC  
CCAGACCTCCCAGTGGGGACGGGGCAGGGCCTGGCCTGGGAAGAGCACAGCTGCAGATCC  
CAGGCCTCTGGCGCCCCCACTCAAGACTACCAAAGCCAGGACACCTCAAGTCTCCAGCCC  
CAATACCCACCCCAATCCCGTATTCTTTTTTTTTTTTTTTTAGACAGGGTCTTGCTCCG  
TTGCCCAGGTTGGAGTGCAGTGGCCCATCAGGGCTCACTGTAACCTCCGACTCCTGGGTT  
CAAGTGACCTCCACCTCAGCCTCTCAAGTAGCTGGGACTACAGGTGCACCACCACACC  
TGGCTAATTTTTGTATTTTTTGTAAAGAGGGGGGTCTCACTGTGTTGCCAGGCTGGTTT  
CGAACTCCTGGGCTCAAGCGGTCCACCTGCCCTCCGCTCCCAAAGTGCTGGGATTGCAGG  
CATGAGCCACTGCACCCAGCCCTGTATTCTTATTCTTCAGATATTTATTTTCTTTTCAC  
TGTTTTTAAATAAAACCAAAGTATTGATAAAAAAAA

26/392

## FIGURE 26

MWRCPLGLLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQDLCCRGRAD  
DCALPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPFPPIQGCMHGGRIYPVLGTYWD  
NCNRCTCQENRQWHGGSRHDQSHQPGQLWLAGWEPQRLLGHDPG

**N-glycosylation site:**

amino acids 78-82, 161-165

**Casein kinase II phosphorylation site:**

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300,  
411-415

**N-myristoylation site:**

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230,  
269-275, 378-384, 442-448

**Amidation site:**

amino acids 26-30, 318-322

**Eukaryotic thiol (cysteine) proteases histidine active site:**

amino acids 398-409

27/392

**FIGURE 27**

CCCACGCGTCCGGCAGGTTTTTCTTCAAGCCAAGAAGGACACGGATTGGCTGAAGGAGAA  
AGTGCAGAGCTTGCAGACACTGGCTGCCAACAACTCTGCGTTGGCCAAAGCCAACAACGA  
CACCTTGGAGGATATGAACAGCCAGCTCAACTCATTACAGGTCAGATGGAGAACATCAC  
CACTATCTCTCAAGCCAACGAGCAGAACCCTGAAAGACCTGCAGGACTTACACAAAGATGC  
AGAGAATAGAACAGCCATCAAGTTCAACCAACTGGAGGAACGCTTCCAGCTCTTTGAGAC  
GGATATTGTGAACATCATTAGCAATATCAGTTACACAGCCCACCACCTGCGGACGCTGAC  
CAGCAATCTAAATGAAGTCAGGACCACTTGCACAGATACCTTACCAAACACACAGATGAT  
CTGACCTCCTTGAATAATACCCTGGCCAACATCCGTTTGGATTCTGTTTCTCTCAGGATG  
CAACAAGATTTGATGAGGTCGAGGTAGACACTGAAGTAGCCAACTTATCAGTGATTATG  
GAAGAAATGAAGCTAGTAGACTCCAAGCATGGTCAGCTCATCAAGAATTTTACAATACTA  
CAAGGTCCACCGGGCCCCAGGGGTCCAAGAGGTGACAGAGGATCCCAGGGACCCCCTGGC  
CCAAGTGGCAACAAGGGACAGAAAGGAGAGAAGGGGGAGCCTGGACCACCTGGCCCTGCG  
GGTGAGAGAGGCCCAATTGGACCAGCTGGTCCCCCGGAGAGCGTGGCGGCAAAGGATCT  
AAAGGCTCCCAGGGCCCCAAAGGCTCCCGTGGTTCCCTGGGAAGCCCGGCCCTCAGGGC  
CCCAGTGGGGACCCAGGCCCCCGGGCCCCACCAGGCAAGAGGGACTCCCCGGCCCTCAG  
GGCCCTCCTGGCTTCCAGGGACTTCAAGGGCACCGTTGGGGAGCCTGGGGTGCTGGACCT  
CGGGGACTGCCAGGCTTGCCTGGGGTACCAGGCATGCCAGGCCCCAAGGGCCCCCCCCGGC  
CCTCCTGGCCCATCAGGAGCGGTGGTGCCCCCTGGCCCTGCAGAATGAGCCAACCCCGGCA  
CCGGAGGACAATAGCTGCCCGCCTCACTGGAAGAACTTCACAGACAAATGCTACTATTTT  
TCAGTTGAGAAAGAAATTTTTGAGGATGCAAAGCTTTTCTGTGAAGACAAGTCTTCACAT  
CTTGTTTTTCATAAACACTAGAGAGGAACAGCAATGGATAAAAAAACAGATGGTAGGGAGA  
GAGAGCCACTGGATCGGCCTCACAGACTCAGAGCGTGAAAATGAATGGAAGTGGCTGGAT  
GGGACATCTCCAGACTACAAAATTGGAAAGCTGGACAGCCGGATAACTGGGGTCATGGC  
CATGGGCCAGGAGAAGACTGTGCTGGGTTGATTTATGCTGGGCAGTGGAACGATTTCCAA  
TGTGAAGACGTCAATAACTTCATTTGCGAAAAAGACAGGGAGACAGTACTGTCATCTGCA  
TTATAACCGGACTGTGATGGGATCACATGAGCAAATTTTCAGCTCTCAAAGGCAAAGGACA  
CTCCTTTCTAATTGCATCACCTTCTCATCAGATTGAAAAAAAAAAGCACTGAAAACCAA  
TTACTGAAAAAAATTTGACAGCTAGTGTTTTTTACCATCCGTCATTACCCAAAGACTTGG  
GAACTAAATGTTCCCCAGGGTGATATGCTGATTTTCATTGTGCACATGGACTGAATCAC  
ATAGATTCTCCTCCGTCAGTAACCGTGCGATTATACAAATTATGTCTTCCAAAGTATGGA  
ACACTCCAATCAGAAAAAGGTTATCATTGGTCGTTGAGTTATGGGAAGAACTTAAGCATA  
TACTGTGTAAACAGTGCCATACATTTCTAAAATCCCAAGTGTAGGAAAAATATGCAGACA  
TACAGATATATAGGCCAACTATTAGTAATAATATGAAATATACTTAAAGAGCTTTTAAAA  
CTTTGTATTTTTGTACAAAAAAA

28/392

**FIGURE 28**

MQQDLMSRLDTEVANLSVIMEEMKLVD SKHGQLIKNFTILQGPPGPRGPRGDRGSQGGP  
GPTGNKGQKGEKGEPPGPAGERGPAGPPGERGGKGSKGSQGPKGSRGSPGKPGPQ  
GPSGDPGPPGPPGKEGLPGPQGPPGFQGLQGTVGEPGVPGPRGLPGLPGVPGMPGPKGPP  
GPPGPSGAVVPLALQNEPTPAPEDNSCPHWKNFTDKCYYFSVEKEIFEDAKLFCEDKSS  
HLVFINTREEQQWIKQMVGRESHWIGLTD SERENWKWLDGTSPDYKNWKAGQPDNWGH  
GHGPGEDCAGLIYAGQW NDFQCEDVNNFICEKDRETVLSSAL

**Signal sequence:**

None

**Transmembrane domain:**

None

**N-glycosylation site:**

16-19, 37-40, 213-216

**Tyrosine kinase phosphorylation site:**

212-220

**N-myristoylation site:**

97-102, 100-105, 148-153, 267-272, 293-298, 310-315

**Cell attachment sequence:**

51-53

**C-type lectin domain signature:**

308-330

**Lectin C-type domain:**

233-330

**Collagen triple helix repeat:**

43-102, 127-186

29/392

**FIGURE 29**

GGACTAATCTGTGGGAGCAGTTTATTCCAGTATCACCCAGGGTGCAGCCACACCAGGACT  
GTGTTGAAGGGTGTTTTTTTCTTTTAAATGTAATACCTCCTCATCTTTTCTTCTTACAC  
AGTGTCTGAGAACATTTACATTATAGATAAGTAGTACATGGTGGATAACTTCTACTTTTA  
GGAGGACTACTCTCTTCTGACAGTCCTAGACTGGTCTTCTACACTAAGACACCATGAAGG  
AGTATGTGCTCCTATTATTCTGGCTTTGTGCTCTGCCAAACCTTCTTTAGCCCTTCAC  
ACATCGCACTGAAGAATATGATGCTGAAGGATATGGAAGACACAGATGATGATGATGATG  
ATGATGATGATGATGATGATGAGGACAACTCTCTTTTCCAACAAGAGAGCCAAGAA  
GCCATTTTTTCCATTTGATCTGTTTCCAATGTGTCCATTTGGATGTCAGTGCTATTTCAC  
GAGTTGTACATTGCTCAGATTTAGGTTTGACCTCAGTCCCAACCAACATTCCATTTGATA  
CTCGAATGCTTGATCTTCAAAACAATAAAATTAAGGAAATCAAAGAAATGATTTTAAAG  
GACTCACTTCACTTTATGGTCTGATCCTGAACAACAAGCTAACGAAGATTCAACCCAA  
AAGCCTTTCTAACCACAAAGAAGTTGCGAAGGCTGTATCTGTCCCACAATCAACTAAGTG  
AAATACCACTTAATCTTCCCAAATCATTAGCAGAACTCAGAATTCATGAAAATAAAGTTA  
AGAAAATACAAAAGGACACATTCAAAGGAATGAATGCTTTACACGTTTTGGAAATGAGTG  
CAAACCTCTTGATAATAATGGGATAGAGCCAGGGGCATTTGAAGGGGTGACGGTGTTCC  
ATATCAGAATTGCAGAAGCAAACTGACCTCAGTTCCTAAAGGCTTACCACCAACTTTAT  
TGGAGCTTCACTTAGATTATAATAAAATTTCAACAGTGGAAGTTGAGGATTTTAAACGAT  
ACAAAGAAGTACAAAGGCTGGGCCTAGGAAACAACAAATCACAGATATCGAAAATGGGA  
GTCTTGCTAACATAACACGTGTGAGAGAAATACATTTGGAAAACAATAAACTAAAAAAA  
TCCCTTCAGGATTACCAGAGTTGAAATACCTCCAGATAATCTTCCTTCATTCTAATTCAA  
TTGCAAGAGTGGGAGTAAATGACTTCTGTCCAACAGTGCCAAAGATGAAGAAATCTTTAT  
ACAGTGCAATAAGTTTATTCAACAACCCGGTGAAATACTGGGAAATGCAACCTGCAACAT  
TTCGTTGTGTTTTGAGCAGAATGAGTGTTTCAGCTTGGGAAGTTTGAATGTAATAATTAG  
TAATTGGTAATGTCCATTTAATATAAGATTCAAAAATCCCTACATTTGGAATACTTGAAC  
TCTATTAATAATGGTAGTATTATATATACAAGCAAATATCTATTCTCAAGTGGTAAGTCC  
ACTGACTTATTTTATGACAAGAAATTTCAACGGAATTTTGCCAAACTATTGATACATAAG  
GGGTGAGAGAAACAAGCATCTATTGCAGTTTCTTTTTTGCGTACAAATGATCTTACATA  
AATCTCATGCTTGACCATTCTTTCTTCATAACAAAAAGTAAGATATTCGGTATTTAAC  
ACTTTGTTATCAAGCACATTTTAAAAGAACTGTACTGTAAATGGAATGCTTGACTTAGC  
AAAATTTGTGCTCTTTCAATTGCTGTTAGAAAAACAGAATTAACAAAGACAGTAATGTGA  
AGAGTGCAATTACACTATTCTTATTCTTTAGTAACCTGGGTAGTACTGTAATATTTTAAAT  
CATCTTAAAGTATGATTTGATATAATCTTATTGAAATTACCTTATCATGTCTTAGAGCCC  
GTCTTTATGTTTTAAACTAATTTCTTAAAAATAAGCCTTCAGTAAATGTTCAATTACCAAC  
TTGATAAATGCTACTCATAAGAGCTGGTTTGGGGCTATAGCATATGCTTTTTTTTTTTTA  
ATTATTACCTGATTTAAAAATCTCTGTAAAAACGTGTAGTGTTTCATAAAATCTGTAAC  
CGCATTTTAAATGATCCGCTATTATAAGCTTTTAAATAGCATGAAAATTGTTAGGCTATATA  
ACATTGCCACTTCAACTCTAAGGAATATTTTGGAGATATCCCTTTGGAAGACCTTGCTTG  
GAAGAGCCTGGACACTAACAATTCTACACCAAATTGTCTCTTCAAATACGTATGGACTGG  
ATAACTCTGAGAAACACATCTAGTATAACTGAATAAGCAGAGCATCAAATTAACAGACA  
GAAACCGAAAGCTCTATATAAATGCTCAGAGTTCTTTATGTATTTCTTATTGGCATTCAA  
CATATGTAAATCAGAAAACAGGGAAATTTTCATTAAAAATATTGGTTTGAAAT

30/392

**FIGURE 30**

MKEYVLLLFLALCSAKPFFSPSHIALKNMMLKDMEDTDDDDDDDDDDDDDEDNSLFPTRE  
PRSHFFPFDLFPMCPFGCQCYSRVVHCSDLGLTSVPTNIPFDTRMLDLQNNKIKEIKEND  
FKGLTSLYGLILNNNKLTKIHPKAFLTTKKLRRLYLSHNQLSEIPLNLPKSLAELRIHEN  
KVKKIQKDTFKGMNALHVLEMSANPLDNNGIEPGAEGVTVFHIRIAEAKLTSVPKGLPP  
TLELHLDYNKISTVELEDFKRYKELQRLGLGNNKITDIENGLANIPRVREIHLENNKL  
KKIPSGLPPELKYLQIIFLHSNSIARVGVNDFCPTVPKMKKSLYSAISLFNNPVKYWEMQP  
ATFRCVLSRMSVQLGNFGM

**Signal sequence:**  
amino acids 1-15

**N-glycosylation site:**  
amino acids 281-285

**N-myristoylation sites:**  
amino acids 129-135, 210-216, 214-220, 237-243, 270-276,  
282-288

**Leucine zipper pattern:**  
amino acids 154-176

31/392

## FIGURE 31

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATA  
GTTCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGC  
CATCTGAGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTG  
GTGTTGCTTCTCACTTCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCA  
GAAGAGCTTTCCATCCAGGTGTCATGCAGAATTATGGGGATCACCCCTTGTGAGCAAAAAG  
GCGAACCAGCAGCTGAATTTACAGAAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGT  
TTGGCCGGCAAGGACCAAGTTGAAACAGCCTTGAAAGCTAGCTTTGAAACTTGCAGCTAT  
GGCTGGGTGGAGATGGATTTCGTGGTCATCTCTAGGATTAGCCCAAACCCCAAGTGTGGG  
AAAAATGGGGTGGGTGTCCTGATTTGGAAGGTTCCAGTGAGCCGACAGTTTGCAGCCTAT  
TGTTACAACCTCATCTGATACTTGGACTAACTCGTGCAATTCCAGAAATTATCACCACCAA  
GATCCCATATTCAACACTCAAACTGCAACACAAACACAGAATTTATTGTCACTGACAGT  
ACCTACTCGGTGGCATCCCCTTACTCTACAATACCTGCCCTACTACTCCTCCTGCT  
CCAGCTTCCACTTCTATTCCACGGAGAAAAAATTGATTTGTGTACAGAAGTTTTATG  
GAACTAGCACCATGTCTACAGAACTGAACCATTGTTGAAAATAAAGCAGCATTCAAG  
AATGAAGCTGCTGGGTTTGGAGGTGTCCACCGGCTCTGCTAGTGCTTGCTCTCCTCTTC  
TTTGGTGCTGCAGCTGGTCTTGGATTTTGCTATGTCAAAAGGTATGTGAAGGCCTTCCCT  
TTTACAAACAAGAATCAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAAG  
GCCAATGATAGCAACCCTAATGAGGAATCAAAGAAAAGTGAATAAAACCCAGAAGAGTCC  
AAGAGTCCAAGCAAACTACCGTGCATGCCTGGAAGCTGAAGTTTAGATGAGACAGAAA  
TGAGGAGACACACCTGAGGCTGGTTTCTTTCATGCTCCTTACCCTGCCCCAGCTGGGGAA  
ATCAAAGGGCCAAAGAACCAAGAAGAAAGTCCACCCTTGGTTCTTAAGTGAATCAGC  
TCAGGACTGCCATTGGACTATGGAGTGCACCAAGAGAATGCCCTTCTCCTTATTGTAAC  
CCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCACGGCCTTTCTAGCCTGGCTAT  
GTCCTAATAATATCCCACTGGGAGAAAGGAGTTTTGCAAAGTGCAAGGACCTAAACATC  
TCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGGTGGGTTGAAAGC  
CAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGACCCTTTCTTCA  
GCTCTGAAAGAGAAACACGTATCCACCTGACATGTCTTCTGAGCCCGGTAAGAGCAAA  
AGAATGGCAGAAAAGTTTAGCCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAA  
TCTCTGTAAAGCTAAAATAAAGAAATAGAACAAAGGCTGAGGATACGACAGTACACTGTCA  
GCAGGGACTGTAAACACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGAAT  
CACTGTTTAGAACACACACACTTACTTTTTCTGGTCTCTACCCTGCTGATATTTTCTCT  
AGGAAATATACTTTTACAAGTAACAAAAATAAAAACTCTTATAAATTTCTATTTTTATCT  
GAGTTACAGAAATGATTACTAAGGAAGATTACTCAGTAATTTGTTTAAAAAGTAATAAAA  
TTCAACAAACATTTGCTGAATAGCTACTATATGTCAAGTGCTGTGCAAGGTATTACACTC  
TGTAATTGAATATTATTCCCTCAAAAATTGCACATAGTAGAACGCTATCTGGGAAGCTAT  
TTTTTTCAGTTTTGATATTTCTAGCTTATCTACTTCCAACTAATTTTTATTTTGGCTGA  
GACTAATCTTATTCATTTTCTCTAATATGGCAACCATTATAACCTTAATTTATTATTAAC  
ATACCTAAGAAGTACATTGTTACCTCTATATACCAAGCACATTTTAAAGTGCCATTAA  
CAAATGTATCACTAGCCCTCCTTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATT  
TGTGACAAAAAATTAAAGCATTTAGAAAACCT

32/392

**FIGURE 32**

MARCFSLVLLLLTSIWTTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQQLNFTEAKEA  
CRLGLSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLIWKVPV  
SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA  
PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAGFGGVPTAL  
LVLALLFFGAAAGLGFCYVKRYVKAFFFTNKNQOKEMIETKVVKEEKANDSNPNEESKKT  
DKNPEESKSPSKTTVRCLEAEV

**Signal sequence:**  
amino acids 1-16

**Transmembrane domain:**  
amino acids 235-254

**N-glycosylation site:**  
amino acids 53-57, 130-134, 289-293

**Casein kinase II phosphorylation site:**  
amino acids 145-149, 214-218

**Tyrosine kinase phosphorylation site:**  
amino acids 79-88

**N-myristoylation site:**  
amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259



## FIGURE 33

[illegible]

34/392

TCTACTTATGTTGGACACTTGGCAGAAGGACCGTGCCCGGCGGCCTCATTTTGACCAGCT  
GGTGGCTGCATTTGACAAGATGATCCGCAAGCCAGATACCCTGCAGGCTGGCGGGGACCC  
AGGGGAAAGGCCTTCCCAGGCCCTTCTGACCCCTGTGGCCCTGGACTTTCCTTGTCTGGA  
CTCACCCCAGGCCTGGCTTTTCAGCCATTGGACTGGAGTGCTACCAGGACAACCTTCTCCAA  
GTTTGGCCTCTGTACCTTCAGTGATGTGGCTCAGCTCAGCCTAGAAGACCTGCCTGCCCT  
GGGCATCACCTTGGCTGGCCACCAGAAGAAGCTGCTGCACCACATCCAGCTCCTTCAGCA  
ACACCTGAGGCAGCAGGGCTCAGTGGAGGTCTGAGAATGACGATACCCGTGACTCAGCCC  
TGGACACTGGTCCGAGAAGGGACATGTGGGACGTGAGCCGGGCTCCAACAGCCTCTGTGA  
GAGATGCCCCACACCAAACCAACCCTCCGATGGCTGCATTCCCTGGTCCCTCCGCTTTTC  
CACCAGCCCCCTCCTCATTAAGGGAAAGAAGGGAATTTGCAAAAAAAAAAAAAAAAAAAAA  
AAAAAAA

35/392

**FIGURE 34**

MATEGAAQLGNRVAGMVCSLWVLLLVSSVLALAEVLLD TTGETSEIGWLTYP PGGWDEVS  
VLDDQRRRLTRTFEACHVAGAPPGTGQDNWLQTHFVERRGAQRAHIRLHFSVRACSSSLGVS  
GGTCRETFTLYYRQAEEPDSPDSVSSWHLKRWTKVDTIAADESFPSSSSSSSSSSSSSSAAW  
AVGPHGAGQRAGLQLNVKERSFGPLTQRGFYVAFQDTGACLALVAVRLFSYTCPAVLRSF  
ASFPETQASGAGGASLVA AVGTCVAHAEEPEEDGVGGQAGGSPRLHCN GEGKWMVAVGGC  
RCQPGYQPARGDKACQACPRGLYKASAGNAPCSPCPARSHAPNPAAPVPCPCLEGFYRASS  
DPPEAPCTGPPSAPQELWFEVQGSALMLHWRLPRELGGRGDLLFN VVCKECEGRQEPASG  
GGGTCHRCRDEVHFDPRQRGLTESRVLVGG LRAHVPYIILEVQAVNGVSELSPDPPQAAAI  
NVSTSHEVPSAVPVVHQVSRASNSITVSWPQPDQTNGNILDYQLRYDDQAEDESHSFTLT  
SETNTATVTQLSPGHIYGFQVRARTAAGHPYGGKVYFQTL PQGELSSQLPERLSLVIGS  
TLGALAFLLLLAAITVLAVVFQRKRGTGYTEQLQQYSSPGLGVKYYIDPSTYEDPCQAIR  
ELAREVDPAYIKIEEVIGTGSFGEVRQGR LQPRGRREQTVAIQALWAGGAESLQMTFLGR  
AAVLGQFQHPNILRLLEGVVT KSRPLMVLTEFMELGPLDSFLRQREGQFSSQLVAMQRGV  
AAAMQYLSSFAFVHRSLSAHSVLVNSHLVCKVARLGHS PQGPSCLLRWAAPEVIAHGKHT  
HVGSDDELWRTALLGHE

**signal sequence:**

Amino acids 1-31

**Transmembrane domains:**

Amino acids 217-234;598-618

**N-glycosylation site:**

Amino acids 481-485

**Glycosaminoglycan attachment sites:**

Amino acids 249-253;419-423

**cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

Amino acids 66-70;150-154;624-628

**Tyrosine kinase phosphorylation sites:**

Amino acids 644-673;664-671

**N-myristoylation sites:**

Amino acids 10-16;15-21;79-85;99-105;118-124;188-194;  
192-198;218-224;250-256;261-267;275-281;276-282;298-304;321-  
327;328-334;420-426;421-427;440-446;449-455;599-605;626-632;  
708-714;766-772;779-785

**Amidation site:**

Amino acids 693-697

**Cell attachment sequences:**

Amino acids 310-313;399-402

36/392

**FIGURE 35**

GGGGTCTCCCTCAGGGCCGGGAGGCACAGCGGTCCCTGCTTGCTGAAGGGCTGGATGTAC  
GCATCCGCAGGTTCCCGCGGACTTGGGGGCGCCCGCTGAGCCCCGGCGCCCGCAGAAGAC  
TTGTGTTTGCCTCCTGCAGCCTCAACCCGGAGGGCAGCGAGGGCCTACCACCATGATCAC  
TGGTGTGTTTTCAGCATGCGCTTGTGGACCCAGTGGGCGTCCTGACCTCGCTGGCGTACTG  
CCTGCACCAGCGGCGGGTGGCCCTGGCCGAGCTGCAGGAGGCCGATGGCCAGTGTCCGGT  
CGACCGCAGCCTGCTGAAGTTGAAAATGGTGCAGGTCGTGTTTCGACACGGGGCTCGGAG  
TCCTCTCAAGCCGCTCCCGCTGGAGGAGCAGGTAGAGTGGAACCCCGAGCTATTAGAGGT  
CCCACCCCAAACCTCAGTTTGATTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATA  
TTCTCCTTACGACTCTCAATACCATGAGACCACCCCTGAAGGGGGGCATGTTTGCTGGGCA  
GCTGACCAAGGTGGGCATGCAGCAAATGTTTGCCCTTGGGAGAGAGACTGAGGAAGAATA  
TGTGGAAGACATTCCCTTTCTTTCACCAACCTTCAACCCACAGGAGGTCTTTATTTCGTTT  
CACTAACATTTTTTCGGAATCTGGAGTCCACCCGTTGTTTGCTGGCTGGGCTTTTCCAGTG  
TCAGAAAGAAGGACCCATCATCATCCACACTGATGAAGCAGATTCAGAAGTCTTGTATCC  
CAACTACCAAAGCTGCTGGAGCCTGAGGCAGAGAACCAGAGGCCGGAGGCAGACTGCCCTC  
TTTACAGCCAGGAATCTCAGAGGATTTGAAAAAGGTGAAGGACAGGATGGGCATTGACAG  
TAGTGATAAAGTGGACTTCTTCATCCTCCTGGACAACGTGGCTGCCGAGCAGGCACACAA  
CCTCCCAAGCTGCCCCATGCTGAAGAGATTTGCACGGATGATCGAACAGAGAGCTGTGGA  
CACATCCTTGACATACTGCCCAAGGAAGACAGGGAAAGTCTTCAGATGGCAGTAGGCCC  
ATTCTCCACATCCTAGAGAGCAACCTGCTGAAAGCCATGGACTCTGCCACTGCCCCGA  
CAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATGTGACCTTCATACCGCTCTTAAT  
GACCCCTGGGGATTTTGTACCACAAATGGCCACCGTTTGCTGTTGACCTGACCATGGAAC  
TTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTATTACCACGGGAAGGAGCA  
GGTGCCGAGAGGTTGCCCTGATGGGCTCTGCCCCGCTGGACATGTTCTTGAATGCCATGTC  
AGTTTATACCTTAAGCCAGAAAAATACCATGCACTCTGCTCTCAAACTCAGGTGATGGA  
AGTTGGAAATGAAGAGTAACTGATTTATAAAAGCAGGATGTGTTGATTTTAAATAAAGT  
GCCTTTATACAATG

37/392

**FIGURE 36**

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKVMQVVFRRHG  
ARSPLKPLPLEEQVEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMF  
AGQLTKVGMQOMFALGERLRKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGL  
FQCQKEGPIIIHTDEADSEVLYPNYQSCWSLRQRTGRRRQTASLQPGISEDLKKVKDRMG  
IDSSDKVDFFILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESLQMA  
VGPFLHILESNNLLKAMDSATAPDKIRKLYLYAAHDVTFIPLMLTGFIDHKWPPFAVDLT  
MELYQHLESKEWFVQLYYHGKEQVPRGCPDGLCPLDMFLNAMS VYTLSPEKYHALCSQTQ  
VMEVGNEE

**Signal sequence:**

amino acids 1-23

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

amino acids 218-222

**Casein kinase II phosphorylation site:**

amino acids 87-91, 104-108, 320-324

**Tyrosine kinase phosphorylation site:**

amino acids 280-288

**N-myristoylation site:**

amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

**Amidation site:**

amino acids 216-220

**Leucine zipper pattern:**

amino acids 10-32

**Histidine acid phosphatases phosphohistidine signature:**

amino acids 50-65

38/392

**FIGURE 37**

ACTGCACTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCG  
ACCCACGCGTCCGCGGACGCGTGGGCGGACGCGTGGGCCGGCTACCAGGAAGAGTCTGCC  
GAAGGTGAAGGCCATGGACTTCATCACCTCCACAGCCATCCTGCCCCCTGCTGTTCCGGCTG  
CCTGGGCGTCTTCGGCCTCTTCCGGCTGCTGCAGTGGGTGCGCGGGAAGGCCTACCTGCG  
GAATGCTGTGGTGGTGATCACAGGCGCCACCTCAGGGCTGGGCAAAGAATGTGCAAAAGT  
CTTCTATGCTGCGGGTGCTAAACTGGTGCTCTGTGGCCGGAATGGTGGGGCCCTAGAAGA  
GCTCATCAGAGAACTTACCGCTTCTCATGCCACCAAGGTGCAGACACACAAGCCTTACTT  
GGTGACCTTCGACCTCACAGACTCTGGGGCCATAGTTGCAGCAGCAGCTGAGATCCTGCA  
GTGCTTTGGCTATGTGACATACTTGTCAACAATGCTGGGATCAGCTACCGTGGTACCAT  
CATGGACACCACAGTGGATGTGGACAAGAGGGTCATGGAGACAACTACTTTGGCCCACT  
TGCTCTAACGAAAGCACTCCTGCCCTCCATGATCAAGAGGAGGCAAGGCCACATTGTCGC  
CATCAGCAGCATCCAGGGCAAGATGAGCATTCCCTTTTCGATCAGCATATGCAGCCTCCAA  
GCACGCAACCCAGGCCTTTCTTTGACTGTCTGCGTGCCGAGATGGAACAGTATGAAATTGA  
GGTGACCGTCATCAGCCCCGGCTACATCCACACCAACCTCTCTGTAAATGCCATCACCGC  
GGATGGATCTAGGTATGGAGTTATGGACACCACCACAGCCCAGGGCCGAAGCCCTGTGGA  
GGTGGCCCAGGATGTTCTTGCTGCTGTGGGGAAGAAGAAGAAAGATGTGATCCTGGCTGA  
CTTACTGCCTTCCTTGGCTGTTTATCTTCGAACTCTGGCTCCTGGGCTCTTCTTCAGCCT  
CATGGCCTCCAGGGCCAGAAAAGAGCGGAAATCCAAGAACTCCTAGTACTCTGACCAGCC  
AGGGCCAGGGCAGAGAAGCAGCACTCTTAGGCTTGCTTACTCTACAAGGGACAGTTGCAT  
TTGTTGAGACTTTAATGGAGATTTGTCTCACAAGTGGGAAAGACTGAAGAAACACATCTC  
GTGCAGATCTGCTGGCAGAGGACAATCAAAAACGACAACAAGCTTCTTCCCAGGGTGAGG  
GGAAACACTTAAGGAATAAATATGGAGCTGGGGTTTAACTAACTAACTAGAAATAAACA  
TCTCAAACAGTAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAG  
CTTGGCCGCCATGGCCCACTTGTTTATTGCAGCTTATAATGGTTAC

39/392

**FIGURE 38**

MDFITSTAILPLLFGCLGVFGLFRLLQWVRGKAYLRNAVVVITGATSGLGKECAKVFYAA  
GAKLVLCGRNGGALEELIRELTASHATKVQTHKPYLVTFDLTDGAIVAAAAEILQCFGY  
VDILVNNAGISYRGTIMDTTVDVDKRVMETNYFGPVALTKALLPSMIKRRQGHIVAIS  
QGKMSIPFRSAYAASKHATQAFFDCLRAEMEQYEIEVTVISPGYIHTNLSVNAITADGSR  
YGVMDTTTAQGRSPVEVAQDVLA AVGKKKKDVILADLLPSLAVYLRTLAPGLFFSLMASR  
ARKERKSKNS

**Signal sequence:**  
amino acids 1-21

**Transmembrane domain:**  
amino acids 104-120, 278-292

**N-glycosylation site:**  
amino acids 228-232

**Glycosaminoglycan attachment site:**  
amino acids 47-51

**Casein kinase II phosphorylation site:**  
amino acids 135-139, 139-143, 253-257

**Tyrosine kinase phosphorylation site:**  
amino acids 145-153, 146-153

**N-myristoylation site:**  
amino acids 44-50, 105-111, 238-244, 242-248, 291-297

**Amidation site:**  
amino acids 265-269

**Prokaryotic membrane lipoprotein lipid attachment site:**  
amino acids 6-17

40/392

## FIGURE 39

GCAAGCCAAGGCGCTGTTTGAGAAGGTGAAGAAGTTCCGGACCCATGTGGAGGAGGGGGACATTGT  
GTACCGCCTCTACATGCGGCAGACCATCATCAAGGTGATCAAGTTCATCCTCATCATCTGCTACAC  
CGTCTACTACGTGCACAACATCAAGTTCGACGTGGACTGCACCGTGGACATTGAGAGCCTGACGGG  
CTACCGCACCTACCGCTGTGCCACCCCTGGCCACACTCTTCAAGATCCTGGCGTCTTCTACAT  
CAGCCTAGTCATCTTCTACGGCCTCATCTGCATGTACACACTGTGGTGGATGCTACGGCGCTCCCT  
CAAGAAGTACTCGTTTGAGTCGATCCGTGAGGAGAGCAGCTACAGCGACATCCCCGACGTCAAGAA  
CGACTTCGCCTTCATGTGCACCTCATTGACCAATACGACCCGCTCTACTCCAAGCGCTTCGCCGT  
CTTCCTGTGCGAGGTGAGTGAGAACAAGCTGCGGCAGCTGAACCTCAACAACGAGTGAGCGCTGGA  
CAAGCTCCGGCAGCGGCTCACCAAGAACGCGCAGGACAAGCTGGAGCTGCACCTGTTTCATGCTCAG  
TGGCATCCCTGACACTGTGTTTGACCTGGTGGAGCTGGAGGTCTCAAGCTGGAGCTGATCCCCGA  
CGTGACCATCCCGCCAGCATTGCCAGCTCAGGGCCTCAAGGAGCTGTGGCTCTACCACACAGC  
GGCCAAGATTGAAGCGCCTGCGCTGGCCTTCCTGCGCGAGAACCTGCGGGCGCTGCACATCAAGTT  
CACCGACATCAAGGAGATCCCGCTGTGGATCTATAGCCTGAAGACACTGGAGGAGCTGCACCTGAC  
GGGCAACCTGAGCGCGGAGAACCAACCGCTACATCGTCATCGACGGGCTGCGGGAGCTCAAACGCCT  
CAAGGTGCTGCGGCTCAAGAGCAACCTAAGCAAGCTGCCACAGGTGGTCAAGATGTGGGCGTGCA  
CCTGCAGAAGCTGTCCATCAACAATGAGGGCACCAAGCTCATCGTCTCAACAGCCTCAAGAAGAT  
GGCGAACCTGACTGAGCTGGAGCTGATCCGCTGCGACCTGGAGCGCATCCCCACTCCATCTTCAG  
CCTCCACAACCTGCAGGAGATTGACCTCAAGGACAACAACCTCAAGACCATCGAGGAGATCATCAG  
CTTCCAGCACCTGCACCGCCTCACCTGCCTTAAGCTGTGGTACAACCACATCGCCTACATCCCCAT  
CCAGATCGGCAACCTCACCAACCTGGAGCGCCTCTACCTGAACCGCAACAAGATCGAGAAGATCCC  
CACCCAGCTCTTCTACTGCCGCAAGCTGCGCTACCTGGACCTCAGCCACAACAACCTGACCTTCCT  
CCCTGCCGACATCGGCCTCCTGCAGAACCTCAGAACCTAGCCATCACGGCCAACCGGATCGAGAC  
GCTCCCTCCGGAGCTCTTCCAGTGCCGGAAGCTGCGGGCCTGCACCTGGGCAACAACGTGTGCA  
GTCACTGCCCTCCAGGGTGGGCGAGCTGACCAACCTGACGCAGATCGAGCTGCGGGGCAACCGGCT  
GGAGTGCTGCTGTGGAGCTGGGCGAGTGCCCACTGCTCAAGCGCAGCGGCTTGGTGGTGGAGGA  
GGACCTGTTCAACACACTGCCACCCGAGGTGAAGGAGCGGCTGTGGAGGGCTGACAAGGAGCAGGC  
CTGAGCGAGGCCGGCCAGCACAGCAAGCAGCAGGACCGCTGCCAGTCCTCAGGCCCGGAGGGGC  
AGGCCTAGCTTCTCCAGAACTCCCGGACAGCCAGGACAGCCTCGCGGCTGGGCAGGAGCCTGGGG  
CCGCTTGTGAGTCAGGCCAGAGCGAGAGGACAGTATCTGTGGGGCTGGCCCCCTTTCTCCCTCTGA  
GACTCAGTCCCCAGGGCAAGTGCTTGTGGAGGAGAGCAAGTCTCAAGAGCGCAGTATTTGGATA  
ATCAGGGTCTCCTCCCTGGAGGCCAGCTCTGCCAGGGGCTGAGCTGCCACCAGAGGTCTTGGGA  
CCCTCACTTTAGTTCTTGGTATTTATTTTTCTCCATCTCCACCTCCTTCATCCAGATAACTTATA  
CATTCCCAAGAAAGTTTCAGCCAGATGGAAGGTGTTTCAGGGAAAGGTGGGCTGCCTTTTCCCTTG  
TCCTTATTTAGCGATGCCGCCGGGCATTTAACACCCACCTGGACTTCAGCAGAGTGGTCCGGGGCG  
AACCAGCCATGGGACGGTCACCCAGCAGTGCCGGGCTGGGCTCTGCGGTGCGGTCCACGGGAGAGC  
AGGCCTCCAGCTGGAAAGGCCAGGCCTGGAGCTTGCTCTTCAGTTTTTGTGGCAGTTTTAGTTTT  
TTGTTTTTTTTTTTTTTAATCAAAAAACAATTTTTTTTTAAAAAAGCTTTGAAAATGGATGGTTTT  
GGGTATTAAGAAAGAAAAAATTTAAAAAAGACACTAACGGCCAGTGAGTTGGAGTCTC  
AGGGCAGGGTGGCAGTTTCCCTTGAGCAAAGCAGCCAGACGTTGAAGTGTGTTTTCTTTCCCTGGG  
CGCAGGGTGCAGGGTGTCTTCCGGATCTGGTGTGACCTTGGTCCAGGAGTTCTATTTGTTCCCTGGG  
GAGGGAGGTTTTTTGTTTTGTTTTTGGGTTTTTTGGTGTCTTGTTTTCTTTCTCCTCCATGTGT  
CTTGGCAGGCACCTCATTTCTGTGGCTGTGCGCCAGAGGGAATGTTCTGGAGCTGCCAAGGAGGGAG  
GAGACTCGGGTTGGCTAATCCCGGATGAACGGTGCTCCATTGCGACCTCCCCCTCCTCGTGCCCTGC  
CCTGCCCTCTCCACGCACAGTGTTAAGGAGCCAAGAGGAGCCACTTCGCCCAGACTTTGTTTCCCCA  
CCTCCTGCGGCATGGGTGTGTCCAGTGCCACCGCTGGCCTCCGCTGCTTCCATCAGCCCTGTGCGC  
ACCTGGTCCCTCATGAAGAGCAGACACTTAGAGGCTGGTCCGGAATGGGGAGGTGCGCCCTGGGAG  
GGCAGGCGTTGGTTCCAAGCCGGTTCCCGTCCCTGGCGCTGGAGTGACACAGCCAGTCCGGCAC  
CTGTTGGCTGGAAGCCAACTGCTTTAGATCACTCGGGTCCCACTTAGAAGGTTCCCGCCTTA  
GATCAATCACGTGGACACTAAGGCACGTTTTAGAGTCTCTTGTCTTAATGATTATGTCATCCGT  
TGTCCTGTCATTTGTGTTTTCTGCGTGTGTCTTGGATATAATCCTCAGAAATAATGCACACTAG  
CCTCTGACAACCATGAAGCAAAAATCCGTTACATGTGGGTCTGAACTTGTAGACTCGGTTCACAGTA  
TCAAATAAATCTATAACAGAAAAA



41/392

**FIGURE 40**

MRQTIKVIKFILIIICYTVYYVHNIFDVEDCTVDIESLTGYRTYRCAHPLATLFKILASF  
YISLVIFYGLICMYTLWWMRLRRSLKKYSFESIREESSYSDIPDVKNDFAFMLHLIDQYDP  
LYSKRFAVFLSEVSENKLRQLNLNNEWTLDKLRQLTKNAQDKLELHFLMLSGIPDTVFD  
LVELEVLELIPDVTIPPSIAQLTGLKELWLYHTAAKIEAPALAFLENLRLALHIKFTD  
IKEIPLWIYSLKTEELHLTGNLSAENNRYIVIDGLRELKRLKVLRLKSNLSKLPQVVTD  
VGVHLQKLSINNEGTKLIVLNSLKKMANLTELELIRCDLERIPHSIFSLHNLQEIDLKDN  
NLKTIEEIIISFQHLHRLTCLKLWYNHIAIYIPIQIGNLTNLERLYLNRNKIEKIPTQLFYC  
RKLRYLDLSHNNLTFLPADIGLLQNLQNLAITANRIETLPPELFQCRKLRLHGLGNNVLQ  
SLPSRVGELTNLTQIELRGNRLECLPVELGECPLLKRSGLVVEEDLFNTLPPEVKERLWR  
ADKEQA

**Transmembrane domain:**

amino acids 51-75 (type II)

**N-glycosylation site:**amino acids 262-266, 290-294, 328-332, 396-400, 432-436,  
491-495**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

amino acids 85-89

**Casein kinase II phosphorylation site:**amino acids 91-95, 97-101, 177-181, 253-257, 330-334, 364-368,  
398-402, 493-497**N-myristoylation site:**

amino acids 173-179, 261-267, 395-401, 441-447

42/392

## FIGURE 41

GGGGGAGAAGGCGGCCGAGCCCCAGCTCTCCGAGCACCGGGTCGGAAGCCGCGACCCGAG  
CCGCGCAGGAAGCTGGGACCGGAACCTCGGCGGACCCGGCCCCACCCAACTCACCTGCGC  
AGGTCACCAGCACCCCTCGGAACCCAGAGGCCCGCGCTCTGAAGGTGACCCCCCTGGGGAG  
GAAGGCGATGGCCCCCTGCGAGGACGATGGCCCCGCGCCCGCTCGCCCCGGCCGGCATCCC  
TGCCGTGCGCCTTGTTGGCTTCTGTGCACGCTCGGCCTCCAGGGCACCCAGGCCGGGCCACC  
GCCCCGCCCCCTGGGCTGCCCCGCGGAGCCGACTGCCTGAACAGCTTTACCGCCGGGGT  
GCCTGGCTTCGTGCTGGACACCAACGCCTCGGTGAGCAACGGAGCTACCTTCCTGGAGTC  
CCCCACCGTGCGCCGGGGCTGGGACTGCGTGCAGCGCTGCTGCACCACCCAGAACTGCAA  
CTTGCGCTAGTGGAGCTGCAGCCCGACCGCGGGGAGGACGCCATCGCCGCTGCTTCCT  
CATCAACTGCCTCTACGAGCAGAACTTCGTGTGCAAGTTCGCGCCAGGGAGGGCTTCAT  
CAACTACCTCACGAGGGAAGTGTACCGCTCCTACCGCCAGCTGCGGACCCAGGGCTTTGG  
AGGGTCTGGGATCCCCAAGGCCTGGGCAGGCATAGACTTGAAGGTACAACCCAGGAACC  
CCTGGTGCTGAAGGATGTGAAAAACACAGATTGGCGCCTACTGCGGGGTGACACGGATGT  
CAGGGTAGAGAGGAAAGACCCAAACCAGGTGGAAGTGTGGGGACTCAAGGAAGGCACCTA  
CCTGTTCCAGCTGACAGTGACTAGCTCAGACCACCCAGAGGACACGGCCAACGTCACAGT  
CACTGTGCTGTCCACCAAGCAGACAGAAGACTACTGCCTCGCATCCAACAAGGTGGGTGCG  
CTGCCGGGGCTCTTTCCACGCTGGTACTATGACCCACGGAGCAGATCTGCAAGAGTTT  
CGTTTATGGAGGCTGCTTGGGCAACAAGAACAACCTTTCGGGAAGAAGAGTGCAATTCT  
AGCCTGTGCGGGGTGTGCAAGGTGGGCCTTTGAGAGGCAGCTCTGGGGCTCAGGCGACTTT  
CCCCCAGGGCCCCCTCCATGGAAGGCGCCATCCAGTGTGCTCTGGCACCTGTGACCCAC  
CCAGTTCGCTGCAGCAATGGCTGCTGCATCGACAGTTTCCTGGAGTGTGACGACACCCC  
CAACTGCCCCGACGCTCCGACGAGGCTGCCTGTGAAAAATACACGAGTGGCTTTGACGA  
GCTCCAGCGCATCCATTTCCCCAGTGACAAAGGGCACTGCGTGGACCTGCCAGACACAGG  
ACTCTGCAAGGAGAGCATCCCCGCGTGGTACTACAACCCCTTCAGCGAACACTGCGCCCCG  
CTTTACCTATGGTGGTTGTTATGGCAACAAGAACAACCTTTGAGGAAGAGCAGCAGTGCCT  
CGAGTCTTGTGCGGCATCTCCAAGAAGGATGTGTTTGGCCTGAGGCGGGAAATCCCCAT  
TCCCAGCACAGGCTCTGTGGAGATGGCTGTACAGTGTTCCTGGTCATCTGCATTGTGGT  
GGTGGTAGCCATCTTGGGTACTGCTTCTTCAAGAACCAGAGAAAGGACTTCCACGGACA  
CCACCACCACCCACCCACCCCTGCCAGCTCCACTGTCTCCACTACCGAGGACACGGGA  
GCACCTGGTCTATAACCACACACCCGCCCCCTCTGAGCCTGGGTCTCACCGGCTCTCAC  
CTGGCCCTGCTTCCTGCTTGCCAAGGCAGAGGCCTGGGCTGGGAAAAACTTTGGAACCAG  
ACTCTTGCCCTGTTTCCAGGCCCACTGTGCCTCAGAGACCAGGGCTCCAGCCCCCTTGG  
AGAAGTCTCAGCTAAGCTCACGTCTGAGAAAGCTCAAAGGTTTGAAGGAGCAGAAAAC  
CCTTGGGCCAGAAAGTACCAGACTAGATGGACCTGCCTGCATAGGAGTTTGGAGGAAGTTG  
GAGTTTTGTTTCTCTGTTCAAAGCTGCCTGTCCCTACCCCATGGTGCTAGGAAGAGGAG  
TGGGGTGGTGTGACACCCCTGGAGGCCCAACCCCTGTCTCCTCCGAGCTCCTCTTCATGCT  
GTGCGCCACAGGCTGGGAGGAAGGACTTCCCTGTGTAGTTTGTGCTGTAAAGAGTTGCTT  
TTTGTTTATTTAATGCTGTGGCATGGGTGAAGAGGAGGGGAAGAGGCCTGTTTGGCCTCT  
CTGTCTCTCTTCTCTTCCCCCAAGATTGAGCTCTCTGCCCTTGATCAGCCCCACCCCTG  
GCCTAGACCAGCAGACAGAGCCAGGAGAGGCTCAGCTGCATTCCGAGCCCCCACCCCA  
AGGTTCTCCAACATCACAGCCAGCCACCCACTGGGTAATAAAAGTGGTTTGTGGAAAA  
AAAAAAAAAAAAAAAAAAAAA

43/392

**FIGURE 42**

MAPARTMARARLAPAGIPAVALWLLCTLGLQGTQAGPPPAPPGLPAGADCLNSFTAGVPG  
FVLDTNASVSNGATFLESPTVRRGWDCVRACCTTQNCNLALVELQPDREGDAIAACFLIN  
CLYEQNFVCKFAPREGFINYLTVREYRSYRQLRTQGFGGSGIPKAWAGIDLVQVQPEPLV  
LKDVENTDWRLLRGDTDVRVERKDPNQVELWGLKEGTYLFQLTVTSSDHPEDTANVTVTV  
LSTKQTEDYCLASNKVGRCRGSFPRWYYDPTEQICKSFVYGGCLGNKNNYLREEECILAC  
RGVQGGPLRGSSGAQATFPQGPMERRHPVCSGTCQPTQFRCSNGCCIDSFLECDDTPNC  
PDASDEAAACEKYTSGFDELQRIHFPSDKGHCVDL PDTGLCKESI PRWYYPFSEHCARFT  
YGGCYGNKNNFEEEEQQCLESCRGISKKDVFGLRREIPI PSTGSVEMAVTVFLVICIVVVV  
AILGYCFFKNQRKDFHGHHPPTPASSTVSTTEDTEHLVYNHTTRPL

**signal sequence:**

Amino acids 1-35

**transmembrane domain:**

Amino acids 466-483

**N-glycosylation sites:**

Amino acids 66-70;235-239;523-527

**N-myristoylation sites:**

A        m        i        n        o                                a        c        i        d        s  
29-35;43-49;161-167;212-218;281-287;282-288;285-291;  
310-316;313-319;422-428;423-429;426-432

**Cell attachment sequence:**

Amino acids 193-199

**Pancreatic trypsin inhibitor (Kunitz) family signatures:**

Amino acids 278-298;419-438

44/392

**FIGURE 43**

CCCACGCGTCCGCACCTCGGCCCCGGGCTCCGAAGCGGCTCGGGGGCGCCCTTTCGGTCA  
ACATCGTAGTCCACCCCCTCCCCATCCCCAGCCCCGGGGATTTCAGGCTCGCCAGCGCCC  
AGCCAGGGAGCCGGCCGGGAAGCGCGATGGGGGGCCCCAGCCGCCCTCGCTCCTGCTCCTGC  
TCCTGCTGTTTCGCCTGCTGCTGGGCGCCCCGGCGGGGCCAACCTCTCCCAGGACGACAGCC  
AGCCCTGGACATCTGATGAAACAGTGGTGGCTGGTGGCACCGTGGTGGTCAAGTGCCAAAG  
TGAAAGATCACGAGGACTCATCCCTGCAATGGTCTAACCTGCTCAGCAGACTCTCTACT  
TTGGGGAGAAGAGAGCCCTTCGAGATAATCGAATTCAGCTGGTTACCTCTACGCCCCACG  
AGCTCAGCATCAGCATCAGCAATGTGGCCCTGGCAGACGAGGGCGAGTACACCTGCTCAA  
TCTTCACTATGCCTGTGCGAACTGCCAAGTCCCTCGTCACTGTGCTAGGAATTCACAGA  
AGCCCATCATCATCTGGTTATAAATCTTCATTACGGGAAAAAGACACAGCCACCCTAAACT  
GTCAGTCTTCTGGGAGCAAGCCTGCAGCCCCGGCTCACCTGGAGAAAGGGTGACCAAGAAC  
TCCACGGGAGAACCAACCCGCATACAGGAAGATCCCAATGGTAAACCTTCACTGTCTAGCA  
GCTCGGTGACATTCCAGGTTACCCGGGAGGATGATGGGGCGAGCATCGTGTGCTCTGTGA  
ACCATGAATCTCTAAAGGGAGCTGACAGATCCACCTCTCAACGCATTGAAGTTTTATACA  
CACCAACTGCGATGATTAGGCCAGACCCTCCCCATCCTCGTGAGGGCCAGAAGCTGTTGC  
TACACTGTGAGGGTCGCGGCAATCCAGTCCCCCAGCAGTACCTATGGGAGAAGGAGGGCA  
GTGTGCCACCCCTGAAGATGACCCAGGAGAGTGCCCTGATCTTCCCTTTCCTCAACAAGA  
GTGACAGTGGCACCTACGGCTGCACAGCCACCAGCAACATGGGCAGCTACAAGGCCTACT  
ACACCCTCAATGTTAATGACCCAGTCCGGTGCCCTCCTCCTCCAGCACCTACCACGCCA  
TCATCGGTGGGATCGTGGCTTTCATTGTCTTCTGCTGCTCATCATGCTCATCTTCCTTG  
GCCACTACTTGATCCGGCACAAAGGAACCTACCTGACACATGAGGCAAAAGGCTCCGACG  
ATGCTCCAGACGCGGACACGGCCATCATCAATGCAGAAGGCGGGCAGTCAGGAGGGGACG  
ACAAGAAGGAATATTTATCTAGAGGGCGCTGCCCACTTCTGCGCCCCCAGGGGCCCT  
GTGGGGACTGCTGGGGCCGTACCAACCCGGACTTGTACAGAGCAACCGCAGGGCCGCCCC  
CTCCCGCTTGCTCCCCAGCCCACCCACCCCTGTACAGAATGTCTGCTTTGGGTGCGGT  
TTTGTACTCGGTTTGGAATGGGGAGGGAGGAGGGCGGGGGGAGGGGAGGGTTGCCCTCAG  
CCCTTTCGCTGGCTTCTCTGCATTTGGGTATTATTATTTTGTAAACAATCCCAAATCAA  
ATCTGTCTCCAGGCTGGAGAGGCAGGAGCCCTGGGGTGAGAAAAGCAAAAAACAACAAA  
AAACA

45/392

**FIGURE 44**

MGAPAAASLLLLLLLLFACCWAPGGANLSQDDSQPWTSDETVVAGGTVVVKCQVKDHEDSSL  
QWSNPAQQTLYFGEKRALRDNRIQLVTSTPHELSSISNVALADEGEYTCSTFTMPVRTA  
KSLVTVLGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPTRIQ  
EDPNGKTFTVSSSVTFQVTREDDGASIVCSVNHESLKGADRSTSQRIEVLYTPTAMIRPD  
PPHPREGQKLLHCEGRGNPVPQQYLWEKEGSVPPLKMTQESALIFPFLNKSDSGTYGCT  
ATSNMGSYKAYYTLNVNDPSPVPSSSSTYHAIIGGIVAFIVFLLLIIMLIFLGHYLIRHKG  
TYLTHEAKGSDDAPDADTAIINAEGGQSGGDDKKEYFI

**Signal sequence:**  
amino acids 1-20

**Transmembrane domain:**  
amino acids 331-352

**N-glycosylation site:**  
amino acids 25-29, 290-294

**Casein kinase II phosphorylation site:**  
amino acids 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

**N-myristoylation site:**  
amino acids 2-8, 23-29, 156-162, 218-224, 295-301, 298-304,  
306-310, 334-340, 360-364, 385-389, 386-390

**Prokaryotic membrane lipoprotein lipid attachment site:**  
amino acids 7-18

46/392

**FIGURE 45**

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCTCCCTGTTGAATTTTTTGCACATGG  
AGGACAGCAGCAAAGAGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATT  
TTACCATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTTCTGGACTTCAACAGAACCCCA  
TCCAGTCATTTTGATTTTGCTGTTTATTTTTTTTTTCTTTTCTTTTCCCACCACATTG  
TATTTTATTTCCGTACTTCAGAAATGGGCTACAGACCACAAAGTGGCCCAGCCATGGGG  
CTTTTTTCTGAAGTCTTGGCTTATCATTTCCCTGGGGCTCTACTCACAGGTGTCCAAAC  
TCCTGGCCTGCCCTAGTGTGTGCCGCTGCGACAGGAACTTTGTCTACTGTAATGAGCGAA  
GCTTGACCTCAGTGCCTCTTGGGATCCCGGAGGGCGTAACCGTACTCTACCTCCACAACA  
ACCAAATTAATAATGCTGGATTTCCTGCAGAACTGCACAATGTACAGTCGGTGCACACGG  
TCTACCTGTATGGCAACCAACTGGACGAATTCCTCATGAACCTTCCCAAGAATGTCAGAG  
TTCTCCATTTGCAGGAAAACAATATTAGACCATTTTACGGGCTGCTCTTGGCCAGCTCT  
TGAAGCTTGAAGAGCTGCACCTGGATGACAACTCCATATCCACAGTGGGGGTGGAAGACG  
GGGCCTTCCGGGAGGCTATTAGCCTCAAATTGTTGTTTTTGTCTAAGAATCACCTGAGCA  
GTGTGCCTGTTGGGCTTCTGTGGACTTGCAAGAGCTGAGAGTGGATGAAAATCGAATTG  
CTGTCATATCCGACATGGCCTTCCAGAATCTCACGAGCTTGGAGCGTCTTATTGTGGACG  
GGAACCTCCTGACCAACAAGGGTATCGCCGAGGGCACCTTCAGCCATCTCACCAAGCTCA  
AGGAATTTTCAATTGTACGTAATTCGCTGTCCACCCCTCCTCCCGATCTCCCAGGTACGC  
ATCTGATCAGGCTCTATTTGCAGGACAACCAGATAAACCACATTCCTTTGACAGCCTTCT  
CAAATCTGCGTAAGCTGGAACGGCTGGATATATCCAACAACCAACTGCGGATGCTGACTC  
AAGGGGTTTTTGTATAATCTCTCCAACCTGAAGCAGCTCACTGCTCGGAATAACCCCTGGT  
TTTGTGACTGCAGTATTAAATGGGTACAGAATGGCTCAAATATATCCCTTCATCTCTCA  
ACGTGCGGGGTTTCATGTGCCAAGGTCCTGAACAAGTCCGGGGGATGGCCGTCAGGGAAT  
TAAATATGAATCTTTTGTCTGTCCCACCACGACCCCCGGCCTGCCTCTCTTACCCCCAG  
CCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACCCTCTCTATTCCAAACCCCTAGCA  
GAAGCTACACGCCTCCAACCTCCTACCACATCGAACTTCCCACGATTCCTGACTGGGATG  
GCAGAGAAAGAGTGACCCACCTATTTCTGAACGGATCCAGCTCTCTATCCATTTTGTGA  
ATGATACTTCCATTCAAGTCAGCTGGCTCTCTCTCTTACCGTGATGGCATACAAACCTCA  
CATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGCATCGTTCAGGAGCGCATAGTCAGCG  
GTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATCGGATTTGTT  
TAGTGCCACTGGATGCTTTTAACTACCGCGCGGTAGAAGACACCATTGTTTCAGAGGCCA  
CCACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGA  
CGTCCCACAGCATGGGCTCCCCCTTCTGCTGGCGGGCTTGATCGGGGGCGCGGTGATAT  
TTGTGCTGGTGGTCTTGCTCAGCGTCTTTTGTCTGGCATATGCACAAAAAGGGGCGCTACA  
CCTCCCAGAAGTGGAATAACAACCGGGGCGGGCGGAAAGATGATTATTGCGAGGCAGGCA  
CCAAGAAGGACAACCTCATCTGGAGATGACAGAAACCAGTTTTTCAGATCGTCTCCTTAA  
ATAACGATCAACTCCTTAAAGGAGATTTAGACTGCAGCCCATTTACACCCCCAAATGGGG  
GCATTAATTACACAGACTGCCATATCCCCAACAACATGCGATACTGCAACAGCAGCGTGC  
CAGACCTGGAGCACTGCCATACGTGACAGCCAGAGGCCAGCGTTATCAAGGCGGACAAT  
TAGACTCTTGAGAACACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAA  
TGTTACACAGATGCATTTGTGCATTTGAATACTCTGTAATTTATACGGTGTACTATATAA  
TGGGATTTAAAAAAGTGCTATCTTTTCTATTTCAAGTTAATTACAAACAGTTTTGTAAC  
TCTTTGCTTTTTTAAATCTT

47/392

**FIGURE 46**

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLG  
IPEGVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENN  
IQTISRALAQLLKLEELHLLDDNSISTVGVEDGAFREAI SLKLLFLSKNHLSSVPVGLPV  
DLQELRVDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRN  
SLSHPPPDLPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLS  
NLKQLTARNNPWFCDCSIKWVTEWLKYIPSSLNVRGFMCGPEQVRGMAVRELNMMNLLSC  
PTTTPGLPLFTPAPSTASPTTQPPTLSIPNPSRSYTPPTPTTSKLPTIPDWDGRERVTPP  
ISERIQLSIHFNVDTSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSL  
VNLEPRSTYRICLVPLDAFNRYRAVEDTICSEATTHASYLNNGSNTASSHEQTTSHSMGSP  
FLLAGLIGGAVIFVLVVLVSVFCWHMHKKGRYTSQKWKNRGRRKDDYCEAGTKKDNSIL  
EMTETSFQIVSLNNDQLLKGD FRLQPIYTPNGGINYTDCHIPNNMRYCNSSVPDLEHCHT

**Signal peptide:**  
amino acids 1-42

**Transmembrane domain:**  
amino acids 542-561

**N-glycosylation site:**  
amino acids 202-206, 298-302, 433-437, 521-525, 635-639,  
649-653

**Casein kinase II phosphorylation site:**  
amino acids 204-208, 407-411, 527-531, 593-597, 598-602,  
651-655

**Tyrosine kinase phosphorylation site:**  
amino acids 319-328

**N-myristoylation site:**  
amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300,  
522-528, 545-551, 633-639

**Amidation site:**  
amino acids 581-585

**Leucine zipper pattern:**  
amino acids 164-186

**Phospholipase A2 aspartic acid active site:**  
amino acids 39-50

48/392

**FIGURE 47**

GCAGCGAGCGCCGGGTGCGGCCCTGCCGCCGACGGGATGTGACCTTCACCGTCGCTTAGC  
CAGGATGACCGGAGCCCGTGTCTCGCGGCGTCCGCGCCTCGCTTCAGCCTCCCGGGTGCT  
CTGACCGCACGCTCCCGGCTGCTAGGCTCCCCGGCACC GGCCCTCGCCATGCCCGCCACCGC  
CCGGGCCCGCCGCCGCCCTGGGCACTGCGCTTCTGCTGCTCCTGCTGGCTTCCGAGTCTT  
CTCACACTGTGCTGTTGCGGGCGCGTGAGGCGGCGCAGTTTCTGCGGCCCAGGCAGCGCC  
GCGCCTACCAAGTCTTCGAGGAGGCCAAGCAGGGCCACCTGGAACGGGAGTGCGTGAGG  
AGGTGTGCAGCAAAGAGGAGGCCAGAGAGGTGTTGAGAACGACCCCGAGACGGAGTATT  
TCTATCCACGATATCAAGAGTGCATGAGAAAATATGGCAGGCCTGAAGAAAAAACCCAG  
ATTTCCGCCAAATGTGTTTCAAACTTGCCCTGACCAAGTGCACCCCAAACCTTGTGATAAGA  
AGGGTACTCATATCTGCCAAGACCTCATGGGCAACTTCTTCTGCGTGTGCACAGATGGCT  
GGGGAGGCCGGCTCTGTGACAAAGATGTCAATGAGTGTGTCCAGAAGAATGGGGCTGCA  
GCCAGGTCTGCCACAACAAACCAGGAAGCTTCCAATGTGCCTGCCATAGTGGCTTCTCGC  
TTGCATCAGACGGCCAGACCTGCCAAGATATCGATGAATGCACAGACTCAGACACCTGTG  
GGGACGCGCGATGCAAGAACTTGCCAGGCTCCTACTCTTGCCTCTGCGATGAGGGATATA  
CATACAGCTCCAAGGAGAAGACCTGCCAAGATGTGGACGAGTGCCAGCAGGATCGCTGTG  
AGCAGACCTGTGTCAACTCCCCAGGCAGCTATACCTGCCACTGTGATGGGCGAGGGGGCC  
TAAAACTATCCCCAGACATGGATACTTGTGAGGACATCTTACCATGTGTGCCCTTCAGCA  
TGGCCAAGAGCGTGAAGTCCTTGTACCTGGGCCGCGATGTTTACGCGGGACCCCGTGATTA  
GACTACGCTTCAAGAGGCTTCAGCCTACCAGGCTGCTGGCTGAATTTGACTTCCGCACTT  
TTGACCCCTGAAGGAGTCTTCTTCTTCTGCTGGAGGCCGTTTACGACAGCACCTGGATTGTCC  
TGGGCCTAAGAGCTGGGCGGCTTGAGCTGCAGCTTCCGTTACAATGGCGTTGGGCGCATCA  
CCAGCAGCGGGCCAACCATCAACCACGGCATGTGGCAAACTATCTCCGTGGAAGAGCTGG  
AACGTAACCTTGTATCAAGGTCAACAAAGATGCTGTAATGAAGATCGCGGTAGCTGGGG  
AGCTGTTTACAGCTGGAGAGGGGGCTTATCACCTGAATCTCACCGTGGGCGGCATTCCCT  
TCAAGGAGAGTGTGAGCTCGTCCAGCCGATTAACCTCGCCTGGATGGGTGCATGAGGAGTT  
GGAAGTGGCTGAACGGGGAAGACAGCGCCATCCAGGAGACAGTCAAGGCAAACACAAAA  
TGCAGTGCCTTCTGTGACAGAAAGGGGCTCCTTCTTCCCGGGGAATGGATTTGCTACCT  
ACAGGCTCAACTACACCCGAACATCGCTGGATGTGCGCACGGAAACACCTGGGAAGTTA  
AAGTTGTGGCTCGGATCCGCCCTGCCACGGACACGGGGGTGCTGCTGGCGCTGGTGGGGG  
ACGACGATGTGCTCATCTCTGTGGCCCTAGTTCGACTACCACTCTACAAAGAAGCTCAAGA  
AGCAGTTGGTGGTCCTGGCAGTTGAGGATGTTGCCCTGGCACTGATGGAAATCAAGGTGT  
GCGACAGCCAGGAACACACGGTCACTGTCTCCCTGCGGGAGGTTGAGGCCACCTAGAAG  
TGGATGGCACAAAGGGCCAGAGTGAAGTGAGCACTGCCCAGCTGCAGGAGCGACTGGACA  
CACTTAAGACACATCTGCAAGGCTCTGTGCACACCTATGTTGGAGGCCTGCCAGAAGTAT  
CGGTGATTTCTGCACCCGTCCTGCGTTCTACCGCGGATGCATGACTCTGGAGGTAAACG  
GGAAAATCCTGGACCTGGATACGGCCTCGTACAAGCACAGTGACATCACCTCCCACTCCT  
GCCCCCTGTGGAGCATGCCACCCCTTAGACCGAGCTGCAAGAGGGCTCCACACCTAAAG  
ACAAAAATGAAGCAGGGTTTGGACACACAGCACTGGCTCCTCTCGCATGGTCTTGCAACA  
CTGGAGCAGCGTGGACCGCCCTTGTGGTTTTTTTTTTCTTGAGATCTTTCTTTTGCCTTG  
TAACATATCTGTACATAATGGACGGGTGTGCGGTACCGGCTGCTCAGAGAGAGCCACGT  
GACCTGGTGGGAGCTGGCTGGAAGGGGCTGGGCTAGAGGGGCTGGCAGTTTGCAGCAGAA  
CGGATGTGAAGAAAATAATCTCTATTATTTTATTACTACATGCTTCTTTCTGACTCTA  
AAATATGGAAAATAAAATATTTACAGAAACCTTTTAAAAA



49/392

**FIGURE 48**

MPPPPGPAAALGTALLLLLLLASESSHTVLLRAREAAQFLRPRQRRAYQVFEEAKQGHLE  
ECVVEVCSKEEAREVFENDPETEYFYPRYQECMRKYGRPEEKNPDAKCVQNLDPQCTPN  
PCDKKGTHICQDLMGNFVCVCTDGGWGGRLCDKDVNECVQKNGGCSQVCHNKPGSFQCACH  
SGFSLASDGQTCQDIDECTDSDTCGDARCKNLPGSYSCLCDEGYTYSSKEKTCQDVDECQ  
QDRCEQTCVNSPGSYTCHCDGRGGLKLSPDMDTCEDILPCVPFMAKSVKSLYLGRMFSG  
TPVIRLRFKRLQPTRLLAEFDFTFDPEGVLFFAGGRSDSTWIVLGLRAGRLELQLRYNG  
VGRITSSGPTINHGMWQTISVEELERNLVIKVNKDAVMKIAVAGELFQLERGLYHLNLT  
GGIPFKESELVQPINPRLDGCMRSWNWLNGEDSAIQETVKANTKMQCFSVTERGSFFPGN  
GFATYRLNYTRTSLDVGTETTWEVKVVARIRPATDTGVLLALVGDDDVVISVALVDYHST  
KKLKKQLVVLAVEDVALALMEIKVCDSEHTVTVSLREGEATLEVDGTKGQSEVSTAQLQ  
ERLDTLKTHLQGSVHTYVGGLPEVSVISAPVTAFYRGCMTLEVNGKILDLDTASYKHSDI  
TSHSCPPVEHATP

50/392

**FIGURE 49**

CGCCGCGCTCCCGCACCCGCGGCCCGCCACCGCGCCGCTCCCGCATCTGCACCCGCAGC  
CCGGCGGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCA  
GTCGGGGCGGCGGCTGCGGGCGCAGAGCGGAGATGCGAGCGGCTTGGGGCCACCCTGCTGT  
GCCTGCTGCTGGCGGCGGCGGTCCCCACGGCCCCCGCGCCCGCTCCGACGGCGACCTCGG  
CTCCAGTCAAGCCCGGCCCGGCTCTCAGTACCCGCAGGAGGAGGCCACCCTCAATGAGA  
TGTTCGCGAGGTTGAGGAACTGATGGAGGACACGCAGCACAAATTGCGCAGCGCGGTGG  
AAGAGATGGAGGCAGAAGAAGCTGCTGCTAAAGCATCATCAGAAAGTGAACCTGGCAAAC  
TACCTCCCAGCTATCACAATGAGACCAACACAGACACGAAGGTTGGAAATAATACCATCC  
ATGTGCACCGAGAAATTCACAAGATAACCAACAACCAGACTGGACAAATGGTCTTTTCAG  
AGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCCACGAGTGCATCATCG  
ACGAGGACTGTGGGCCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTACACCTGCCAGC  
CATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGCTGTGGAGACCAGCTGT  
GTGTCTGGGGTCACTGCACCAAATGGCCACCAGGGGCAGCAATGGGACCATCTGTGACA  
ACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCCTTCCAGAGAGGCCTGCTGTTCCCTG  
TGTGCACACCCCTGCCCGTGGAGGGCGAGCTTTGCCATGACCCCGCCAGCCGGCTTCTGG  
ACCTCATACCTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGATGCCCTTGTGCCAGTG  
GCCTCCTCTGCCAGCCCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCTGTG  
GGAGCCGTGACCAAGATGGGGAGATCCTGCTGCCCAGAGAGGTCCCGATGAGTATGAAG  
TTGGCAGCTTCATGGAGGAGGTGCCCCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTG  
AAGAGATGGCGCTGGGGGAGCCTGCCGGCTGCCCGCGCTGCACTGCTGGGAGGGGAAGAGA  
TTTAGATCTGGACCAGGCTGTGGGTAGATGTGCAATAGAAATAGCTAATTTATTTCCCCA  
GGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTACATCTTCTTCCCAGTAAGTT  
TCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTTCAGCTCCCCCAGGCTGTTCT  
CCAGGCTTCACAGTCTGGTGTCTGGGAGAGTCAGGCAGGGTTAAACTGCAGGAGCAGTTT  
GCCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAGACAGCCGTTTGT  
TCTACATGGCTTTGATAATTGTTTGAGGGGAGGAGATGGAAACAATGTGGAGTCTCCCTC  
TGATTGGTTTTTGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAACCTGGCAA  
AAATGCAACAAATGAATTTTCCACGCAGTCTTTTCCATGGGCATAGGTAAGCTGTGCCTT  
CAGCTGTTGCAGATGAAATGTTCTGTTACCCCTGCATTACATGTGTTTATTCATCCAGCA  
GTGTTGCTCAGCTCCTACCTCTGTGCCAGGGCAGCATTTTTCATATCCAAGATCAATTCCC  
TCTCTCAGCACAGCCTGGGGAGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAG  
GCTCAGAGACTGCAAGCTGCTTGCCCAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTC  
ATCTGGTTGTGACTCTAAGCTCAGTGCTCTCTCCACTACCCACACCAGCCTTGGTGCCA  
CCAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGATTTTCTTGGAGGCATGCACATCTGGA  
ATTAAGGTCAAACATAATTCTCACATCCCTCTAAAAGTAACTACTGTTAGGAACAGCAGT  
GTTCTCACAGTGTGGGGCAGCCGTCTTCTAATGAAGACAATGATATTGACACTGTCCCT  
CTTTGGCAGTTGCATTAGTAACTTTGAAAGGTATATGACTGAGCGTAGCATACAGGTTAA  
CCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGCAAAAT  
CACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC  
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAA  
TGATGTTTTTCAAGGTGTCATGGACTGTTGCCACCATGTATTTCATCCAGAGTTCTTAAAGTT  
TAAAGTTGCACATGATTGTATAAGCATGCTTTCTTTGAGTTTTTAAATTATGTATAAACAT  
AAGTTGCATTTAGAAATCAAGCATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAA  
AAAAAA

51/392

**FIGURE 50**

MQRLGATLLCLLLAAVPTAPAPAPTATSAPVKPGPALSYQEEATLNEMFREVEELMED  
TQHKLRSAVEEMEAEEAAKASSEVNLANLPPTYHNETNTDTKVGNNTIHVHREIHKITN  
NQTGQMVFSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRMCLTR  
DSECCGDQLCVWGHCTKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGEL  
CHDPASRLLDLITWELEPDGALDRCPASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILL  
PREVPDEYEVGSFMEEVRQELEDLERSLTEEMALGEPAAAAAALLGGEI

**Signal sequence:**

amino acids 1-19

**N-glycosylation site:**

amino acids 96-100, 106-110, 121-125, 204-208

**Casein kinase II phosphorylation site:**amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,  
327-331**N-myristoylation site:**

amino acids 202-208, 217-223

**Amidation site:**

amino acids 140-144

52/392

**FIGURE 51**

GCCTGTTGCTGATGCTGCCGTGCGGTACTTGTCATGGAGCTGGCACTGCGGCGCTCTCCC  
GTCCCGCGGTGGTTGCTGCTGCTGCCGCTGCTGCTGGGCCTGAACGCAGGAGCTGTCATT  
GACTGGCCACAGAGGAGGGCAAGGAAGTATGGGATTATGTGACGGTCCGCAAGGATGCC  
TACATGTTCTGGTGGCTCTATTATGCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCC  
CTGGTCATGTGGCTTCAGGGCGGTCCAGGCGGTTCTAGCACTGGATTTGGAACTTTGAG  
GAAATTGGGCCCCCTTGACAGTGATCTCAAACCACGGAAAACCACCTGGCTCCAGGCTGCC  
AGTCTCCTATTTGTGGATAATCCCGTGGGCACTGGGTTCA GTTATGTGAATGGTAGTGGT  
GCCTATGCCAAGGACCTGGCTATGGTGGCTTCAGACATGATGGTTCCTCTGAAGACCTTC  
TTCAGTTGCCACAAAGAATTCCAGACAGTTCATTCTACATTTTCTCAGAGTCTTATGGA  
GGAAAAATGGCAGCTGGCATTGGTCTAGAGCTTTATAAGGCCATTACAGCGAGGGACCATC  
AAGTGCAACTTTGCGGGGGTTGCCTTGGGTGATTCCCTGGATCTCCCCTGTTGATTCCGGTG  
CTCTCCTGGGGACCTTACCTGTACAGCATGTCTCTTCTCGAAGACAAAGGTCTGGCAGAG  
GTGTCTAAGGTTGCAGAGCAAGTACTGAATGCCGTAAATAAGGGGCTCTACAGAGAGGCC  
ACAGAGCTGTGGGGGAAAGCAGAAATGATCATTGAACAGAACACAGATGGGGTGAACCTTC  
TATAACATCTTAACTAAAAGCACTCCACGTCTACAATGGAGTCGAGTCTAGAATTCACA  
CAGAGCCACCTAGTTTGTCTTTGTGTCAGCGCCACGTGAGACACCTACAACGAGATGCCTTA  
AGCCAGCTCATGAATGGCCCCATCAGAAAGAAGCTCAAAATTATTCTGAGGATCAATCC  
TGGGGAGGCCAGGCTACCAACGTCTTTGTGAACATGGAGGAGGACTTCATGAAGCCAGTC  
ATTAGCATTGTGGACGAGTTGCTGGAGGCAGGGATCAACGTGACGGTGTATAATGGACAG  
CTGGATCTCATCGTAGATACCATGGGTGAGGAGCCCTGGGTGCGGAACTGAAGTGGCCA  
GAACTGCCTAAATTGAGTCAGCTGAAGTGAAGGCCCTGTACAGTGACCCTAAATCTTTG  
GAAACATCTGCTTTTGTCAAGTCCTACAAGAACCTTGCTTTCTACTGGATTCTGAAAGCT  
GGTCATATGGTTCTTCTGACCAAGGGGACATGGCTCTGAAGATGATGAGACTGGTGACT  
CAGCAAGAATAGGATGGATGGGGCTGGAGATGAGCTGGTTTGGCCTTGGGGCACAGAGCT  
GAGCTGAGGCCGCTGAAGCTGTAGGAAGCGCCATTCTTCCCTGTATCTAACTGGGGCTGT  
GATCAAGAAGGTTCTGACCAGCTTCTGCAGAGGATAAAATCATTGTCTCTGGAGGCAATT  
TGGAATTTATTTCTGCTTCTTAAAAAACCTAAGATTTTTTAAAAAATTGATTTGTTTTG  
ATCAAAATAAAGGATGATAATAGATATTAA

53/392

**FIGURE 52**

MELALRRSPVPRWLLLLPLLLGLNAGAVIDWPTEEGKEVWDYVTVRKDAYMFWWLYYATN  
SCKNFSELPLVMWLQGGPGGSSTGFGNFEEIGPLDSLKPRKTTWLQAASLLFVDNPVGT  
GFSYVNGSGAYAKDLAMVASDMMVLLKTFFSCHKEFQTVPFYIFSESYGGKMAAGIGLEL  
YKAIQRGTIKCNFAGVALGDSWISPVDSVLSWGPYLYSMSLLEDKGLAEVSKVAEQVLNA  
VNKGLYREATELWGKAEMIEQNTDGVNFYNILTKSTPTSTMESSLEFTQSHLVCLCQRH  
VRHLQRDALSQLMNGPIRKKLKIIPEDQSWGGQATNVFVNMEEDFMKPVISIVDELLEAG  
INVTVYNGQLDLIVDTMGQEAWVRKLKWPELPKFSQLKWKALYSDPKSLETSAFVKSYKN  
LAFYWILKAGHMVPSDQGDMAKMMRLVTQQE

**Signal sequence:**

amino acids 1-25

**N-glycosylation site:**

amino acids 64-68, 126-130, 362-366

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

amino acids 101-105

**Casein kinase II phosphorylation site:**amino acids 204-208, 220-224, 280-284, 284-288, 351-355,  
449-453**N-myristoylation site:**amino acids 22-28, 76-82, 79-85, 80-86, 119-125, 169-175,  
187-193, 195-201, 331-337, 332-338, 360-366

54/392

**FIGURE 53**

GTCTGTTCCCAGGAGTCCTTCGGCGGCTGTTGTGTCAGTGGCCTGATCGCGATGGGGACA  
AAGGCGCAAGTCGAGAGGAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCC  
CTGGCATTGGGCAGTGTTACAGTGCACCTCTTCTGAACCTGAAGTCAGAATTCCTGAGAAT  
AATCCTGTGAAGTTGTCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAGTGGAAAG  
TTTGACCAAGGAGACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTAT  
GAGGACCGGGTGACCTTCTTGCCAACTGGTATCACCTTCAAGTCCGTGACACGGGAAGAC  
ACTGGGACATACACTTGTATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAG  
GTCAAGCTCATCGTGCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCCTCCTCTGCC  
ACCATTTGGGAACCGGGCAGTGCTGACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAA  
TACACCTGGTTCAAAGATGGGATAGTGATGCCTACGAATCCCAAAGCACCCGTGCCTTC  
AGCAACTCTTCTATGTCTGAATCCCAACAACAGGAGAGCTGGTCTTTGATCCCCTGTCA  
GCCTCTGATACTGGAGAATAACAGCTGTGAGGCACGGAATGGGTATGGGACACCCATGACT  
TCAAATGCTGTGCGCATGGAAGCTGTGGAGCGGAATGTGGGGGTTCATCGTGGCAGCCGTC  
CTTGTAACCCCTGATTCTCCTGGGAATCTTGGTTTTTGGCATCTGGTTTGCCTATAGCCGA  
GGCCACTTTGACAGAACAAAGAAAGGGACTTCGAGTAAGAAGGTGATTTACAGCCAGCCT  
AGTGCCCGAAGTGAAGGAGAATTCAAACAGACCTCGTCATTCTTGGTGTGAGCCTGGTCG  
GCTCACCGCCTATCATCTGCATTTGCCTTACTCAGGTGCTACCGGACTCTGGCCCCTGAT  
GTCTGTAGTTTACAGGATGCCTTATTTGTCTTCTACACCCACAGGGCCCCCTACTTCT  
TCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCCATCCTCCTTCATGCCCTCCCTCCC  
TTTCCTACCACTGCTGAGTGGCCTGGAACCTTGTTTAAAGTGTTTATTCCCCATTTCTTTG  
AGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCCTTCTAAGTAGACAGCAAAAA  
TGGCGGGGGTTCGAGGAATCTGCACTCAACTGCCCACCTGGCTGGCAGGGATCTTTGAAT  
AGGTATCTTGAGCTTGGTTCTGGGCTCTTTCCTTGTGTACTGACGACCAGGGCCAGCTGT  
TCTAGAGCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGGTGATGACACTGGGG  
TCCTTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTGCCACTGGGATCC  
CTCTGCCCTGTCTCCTGAATACAAGCTGACTGACATTGACTGTGTCTGTGGAAAATGGG  
AGCTCTTGTTGTGGAGAGCATAGTAAATTTTTCAGAGAACTTGAAGCCAAAAGGATTTAAA  
ACCGCTGCTCTAAAGAAAAGAAAAGTGGAGGCTGGGCGCAGTGGCTCACGCCTGTAATCC  
CAGAGGCTGAGGCAGGCGGATCACCTGAGGTGGGAGTTCGGGATCAGCCTGACCAACAT  
GGAGAAACCCTACTGGAAATACAAAGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTCCC  
AGCTGCTCAGGAGCCTGGCAACAAGAGCAAACTCCAGCTCAAAAAAAAAAAAAAAAAA

55/392

**FIGURE 54**

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLS CAYSGFSSPRV  
EWKFDQGD TTRLVCYNNKITASYEDRVTF LPTGITFKSVTREDTGT YTCMVSEEGNSYG  
EVKVKLIVLVPPSKPTVNI PSSATIGNRAVLTCSEQDGSP PSEYTWFKDGIVMPTNPKST  
RAFSNSSYVLNPTTGELVFDPLSASDTGEYSCEARNGYGT PMTSNAVRMEAVERNVGVIV  
AAVLVTLILLGILVFGIWFAYS RGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

**Signal sequence:**  
amino acids 1-27

**Transmembrane domain:**  
amino acids 238-255

**N-glycosylation site:**  
amino acids 185-189

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**  
amino acids 270-274

**Casein kinase II phosphorylation site:**  
amino acids 34-38, 82-86, 100-104, 118-122, 152-156,  
154-158, 193-197, 203-207, 287-291

**N-myristoylation site:**  
amino acids 105-111, 116-122, 158-164, 219-225, 237-243,  
256-262

56/392

**FIGURE 55**

GTTGTGTCCTTCAGCAAAACAGTGGATTTAAATCTCCTTGACACAAGCTTGAGAGCAACAC  
AATCTATCAGGAAAGAAAGAAAGAAAAAACCGAACCTGACAAAAAGAAGAAAAAGAAG  
AAGAAAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAAT  
CTTCACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGC  
CACCTTCCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCCTCAG  
GTGCACTATTGACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTA  
TGCTGGGAATGACAAGTGGTGCCTGGATCCTCGCGTGGTCTTCTGAGCAACACCCAAAC  
GCAGTACAGCATCGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTC  
GGTGCAGACAGACAACACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCC  
CAAAATTGTAGAGATTTCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCAC  
CTGCATAGCAACTGGTAGACCAGAGCCTACGGTTACTTGGAGACACATCTCTCCAAAGC  
GGTTGGCTTTGTGAGTGAAGACGAATACTTGGAAATTTCAGGGCATCACCCGGGAGTGC  
AGGGGACTACGAGTGCAGTGCCTCCAATGACGTGGCCGCGCCCGTGGTACGGAGAGTAAA  
GGTCACCGTGAACCTATCCACCATACATTTCAGAAGCCAAGGGTACAGGTGTCCCCGTGGG  
ACAAAAGGGGACACTGCAGTGTGAAGCCTCAGCAGTCCCCTCAGCAGAATTCCAGTGGTA  
CAAGGATGACAAAAGACTGATTGAAGGAAAGAAAGGGGTGAAAGTGGAACAGACCTTT  
CCTCTCAAACTCATCTTCTTCAATGTCTCTGAACATGACTATGGGAACTACACTTGCGT  
GGCCTCCAACAAGCTGGGCCACACCAATGCCAGCATCATGCTATTTGGTCCAGGCGCCGT  
CAGCGAGGTGAGCAACGGCACGTGAGGAGGGCAGGCTGCGTCTGGCTGCTGCCTCTTCT  
GGTCTTGACCTGCTTCTCAAATTTTGATGTGAGTGCCACTTCCCCACCCGGGAAAGGCT  
GCCGCCACCACCACCACCAACACAACAGCAATGGCAACACCGACAGCAACCAATCAGATA  
TATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGAAATTTGAGGGAGGGGAAC  
AAAGAATACTTTGGGGGGAAAAGAGTTTTAAAAAAGAAATTGAAAATTGCCTTGAGATA  
TTTAGGTACAATGGAGTTTTCTTTTCCCAAACGGGAAGAACACAGCACACCCGGCTTGA  
CCCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGGCAGTGTGGGCAAGGGCTCAGCCTC  
TCTGCCCACAGAGTGCCCCACGTGGAACATTCTGGAGCTGGCCATCCCAAATTCATCA  
GTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCCAAGCGTGGCGCTGCGGGCACTTTG  
GTAGACTGTGCCACCACGGCGTGTGTTGTGAAACGTGAAATAAAAAAGAGCAAAAAAAA



57/392

**FIGURE 56**

MKTIQPKMHNSISWAI FTGLAALCLFQGVPRSGDATFPKAMDNVTVRQGESATLRCTID  
NRVTRVAWLNRSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNV DVYDEGPYTCSVQTD  
NHPKTSRVHLIVQVSPKIVEISSDISINEGNNISLT CIATGRPEPTVTWRHISPKAVGFV  
SEDEYLEIQGITREQSGDYEC SASNDVAAPVRRVKVTVNYPPYISEAKGTGVPVGQKGT  
LQCEASAVPSAEFQWKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGN YTCVASNK  
LHTNASIMLFGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLKF

58/392

**FIGURE 57**

GCTGCGCCGGCTGCGGCTGCAGGGGAATCCGCTGTGGTGCGGCTGCCAGGCGCGGCCCT  
ACTCGAGTGGCTGGCGCGGGCGCGCTGCGCTCGGACGGCGCGTGCCAGGGGCGCGGCG  
CCTGCGGGGCGAGGCTCTGGACGCCCTGCGGCCCTGGGACCTGCGCTGCCCTGGGGACGC  
GGCGCAGGAAGAGGAAGAGCTGGAAGAGCGGGCTGTGGCCGGGCCCCGCGCCCCCTCCGCG  
CGGCCCTCCGCGCGGCCCCGGGGAGGAGCGGGCAGTCGCGCCTTGCCCTCGCGCCTGCGT  
GTGCGTCCCCGAGTCCCGGCACAGCAGCTGCGAGGGCTGCGGCCTGCAGGCGGTGCCCCG  
CGGCTTCCCCAGCGACACCCAGCTCCTGGACCTGAGGCGGAACCACTTCCCCCTCGGTGCC  
CCGAGCGGCCTTCCCCGGNCTGGGCCACCTGGTGTGCTGACCTGCAGCACTGCGGCAT  
CGCGGAGCTGGAAGCGGGCGCCCTGGCCGGGCTGGGCCGCTGATCTACCTGTACCTCTC  
CGACAACCAGCTCGCAGGCCTCAGCGCTGCTGCCCTTGAAGGGGCTCCCCGCCCTCGGCTA  
CCTGTACCTAGAACGCAACCGTTTCTGAGGTGCCAGGGGCTGCCNTGCGCGCCCTGCC  
CAGCCTCTTCTCCCTGCACCTGCAGGACAACGCTGTGGACCGCCTGGCACCTGGGGACCT  
GGGGAGAACACGGGCCTTGCGCTGGGTCTACCTGAGTGGAACCGCATCACCGAAGTGTC  
CCTTGGGGCGCTGGGCCAGCTCGGGAGCTGGAGAAGCTGCACCTGGACAGGAATCAGCT  
GCGAGAGGTGCCCACTGGGGCCTTGAGAGGGCTGCCTGCCCTCCTGGAGCTGCAGCTCTC  
GGGCAACCCACTCAGGGCCTTGCGTGACGGAGCCTTCCAGCCTGTGGGCAGGTGCTGCA  
GCACCTCTTCTGAACAGCAGTGGCCTGGAGCAGATTTGTCCTGGGGCCTTTTCAGGCCT  
GGGGCCCGGGCTCCAGAGCCTGCACCTGCAGAAGAACAGCTTCGGGCCCTGCCTGCCCT  
GCCCAGTCTCAGCCAGCTGGAGCTCATCGACCTCAGCAGCAATCCCTTCCCCTGTGACTG  
CCAGCTGCTTCCGCTGCACAGGTGGCTTACTGGGCTGAACCTGCGGGTGGGGGCCACCTG  
CGCCACCCCTCCCAATGCCCGTGGCCAGAGGGTGAAGGCTGCAGCTGCTGTCTTTGAAGA  
CTGCCCCGGGCTGGGCTGCCAGAAAGGCCAAGCGGACACCAGCCTCCAGGCCCAGTGCCAG  
GAGAACCCCCATCAAAGGAAGACAGTGTGGAGCAGATAAGAACATCCTCTTCCCCACATG  
GTACCACACTGTGGAGCCCACCTCGCTGTCATAGGCCCTGCGGCTCTGAAGGATGGCTTTG  
CCCGCTCCCGCTCTGCCCCCTCAAGTGGAACCCAAGCTGGGCTCAGAATCTGTAGAGTGAG  
GCCCCACCAAGGGAAACGACACCCACGGCCTGAGAGCCAGGTGGAGTCCTGCCACTCAGC  
TGCCTGCCTTTGCTCCACCCCTCTCCACCCCTCAAAGAGGTCTCGAGGGGACACTCTGAA  
GGCACCTGGCTCAGAACCACTGCCATCCAAGGAGCGAGGAGTCCAGGGGCTGAGCAAATG  
CAGCGGGGAGGTGGCAGTTCCCCTGCTTCCCGATCCTCATTTTCTGCTTCACTTGACTC  
CTCCAGATAGGAGCTGCTCTCACTGCCCACACTGCTG

59/392

**FIGURE 58**

LRRLRLQGNPLWCGCQARPLLEWLLARARVRSDGACQGPRRLRGEALDALRPWDLRCPGDA  
AQEEEELEERAVAGPRAPPRGPPRGPEERAVAPCPACVCPESRHSSCEGCGLQAVPR  
GFPSDTQLLDLRRNHFPSPRAAFPGGLHVLVSLHLQHCGIAELEAGALAGLGRLIYLYLS  
DNQLAGLSAAALEGAPRLGYLYLERNRFLQVPGAAXRALPSLFSHLQDNAVDR LAPGDL  
GRTRALRWVYLSGNRITEVSLGALGPARELEKLHLDRNQLREVPTGALEGLPALLELQLS  
GNPLRALRDGAFQPVGRSLQHLFLNSSGLEQICPGAFSGLGPGLSHLQKNQLRALPAL  
PSLSQLELIDLSSNPFPCDCQLLPLHRWLTGLNLRVGATCATPPNARGQRVKAAA VFED  
CPGWAARKAKRTPASRPSARRTPIKGRQCGADKNILFPTWYHTVEPTSLS

**Signal sequence:**

None

**Transmembrane domain:**

None

**N-glycosylation site:**

325-328

**Glycosaminoglycan attachment site:**

338-341

**Protein kinase C phosphorylation site:**

438-440

**N-myristoylation site:**

166-171, 186-191, 253-258, 286-291, 335-340, 339-344, 450-455

**Leucine rich repeat N-terminal domain:**

94-123

**Leucine Rich Repeat:**

125-148, 149-172, 173-196, 197-220, 221-244, 245-268, 269-292, 293-316, 318-341, 343-364, 365-386

**Leucine rich repeat C-terminal domain:**

374-422

60/392

## FIGURE 59

CTCCACGGTGTCCAGCGCCAGAAATGCGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTG  
CTCCAGGTTATGAAGCCCTGGAGGGCCAGAGGAAATCAGCGGGTTCGAAGGGGACACT  
GTGTCCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGG  
AAGGGTGGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAG  
GAGACAATGAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTTGTG  
ACCCTGTGGAACCTCACCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGG  
GGCCCCGATGAGTCTTTACTGATCTCTCTGTTCGTCTTTCCAGGACCCTGCTGTCTCTCC  
TCCCCTTCTCCACCTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCT  
CAGCAAACCCAGCCCCCAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCC  
AAGCAGGGGAAGACAGGGGCTGAGGCCCTCCATTGCCAGGGACTTCCCAGTACGGGCAC  
GAAAGGACTTCTCAGTACACAGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGG  
AGCTCCCGCCCCCCCCATGCAGCTGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTC  
AGCAGTGGCAGCTCTAAGCCCAGGGTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTC  
CTGGTGTCTGCTGAGCCTTCTGTGAGCCGAGGCTGATCGCCTTCTGCAGCCACCTGCTC  
CTGTGGAGAAAGGAAGCTCAACAGGCCACGGAGACACAGAGGAACGAGAAGTTCTGGCTC  
TCACGCTTGACTGCGGAGGAAAAGGAAGCCCCCTTCCCAGGCCCTGAGGGGGACGTGATC  
TCGATGCCTCCCCTCCACACATCTGAGGAGGAGCTGGGCTTCTCGAAGTTTGTCTCAGCG  
TAGGGCAGGAGGCCCTCCTGGCCAGGCCAGCAGTGAAGCAGTATGGCTGGCTGGATCAGC  
ACCGATTCCCAGAAAGCTTTCCACCTCAGCCTCAGAGTCCAGCTGCCCGGACTCCAGGGCT  
CTCCCCACCCTCCCAGGCTCTCCTCTTGATGTTCCAGCCTGACCTAGAAGCGTTTGTCTC  
AGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTGGAGACTGGGACATCCCTGAT  
AGGTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCAGCAGGGCCAGACAAGGCT  
CAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACCTCCTGGGCCCTCATGCCAGTGTCTG  
GACCTTGCCTTCTCTCCACTCCAGACCCACCTTGTCTTCCCTCCCTGGCGTCTCTCAGAC  
TTAGTCCCACGGTCTCCTGCATCAGCTGGTGTATGAAGAGGAGCATGCTGGGGTGAGACTG  
GGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCCAGGAAGCCTGTGAAAAACG  
TGATTCTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAGGACTCTGA  
ATTCTAACAATGCCAGTGAAGTGTGCACTTGAGTTTGAGGGCCAGTGGGCCTGATGAAC  
GCTCACACCCCTTCACTTAGAGTGTGATTTGGGCTGTGACGTCTCCACCTGCCCCAAT  
AGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAGTC  
CAGGCCTTTGGTCAGGTCAAGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG  
TTGCCCTTTNCCATTGCCCCCTCCCTGGNCCATGCCTTCTTGCCTTTGGAAAAAATGATGAA  
GAAAACCTTGGCTCCTTCCCTTGTCTGGAAAGGGTTACTTGCCTATGGGTTCTGGTGGCTA  
GAGAGAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACA  
GGGCGGATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTGGGGGTGGTG  
GTAAAGTAGCACAACCTACTATTTTTTTTCTTTTTTCCATTATTATTGTTTTTTAAGACAGA  
ATCTCGTGCTGCTGCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACTCCGCCTCCTGG  
GTTCAAGTGATTCTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACC  
ACACCTGGCTAATTTTTTGTACTTTTAGTAGAGATGGGGTTTACCATTGTTGGCCAGGCTG  
GTCTTGAACCTCCTGACCTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTA  
CAGGCATGAGCCACTGTGTCTGGCCCTATTTCTTTAAAAAGTGAAATTAAGAGTTGTTT  
AGTATGCAAACTTGGAAAGATGGAGGAGAAAAAGAAAAGGAAGAAAAAATGTACCCCA  
TAGTCTCACCAGAGACTATCATTATTTCTGTTTTGTTGTACTTCTTCCACTCTTTTCTTC  
TTCACATAATTTGCCGGTGTCTTTTTTACAGAGCAATTATCTTGTATATACAACCTTTGTA  
TCCTGCCTTTTTCCACCTTATCGTTCCATCACTTTATTCCAGCACTTCTCTGTGTTTTACA  
GACCTTTTTATAAATAAAATGTTTCATCAGCTGCATAAAAAAAAAAAAAA

61/392

**FIGURE 60**

MRLLVLLWGCLLLPGYEALLEGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSR  
CSGTIYAE EEGQETMKGRVSIRDSRQELSLIVTLWNLTLDAGEYWCGVEKRGPDSESLI  
SLFVFPGPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAE  
APPLPGTSQYGHERTSQYTGTSPHPATSPPAGSSRPPMQLDSTSAEDTSPALSSGSSKPR  
VSIPMVRILAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAEEK  
EAPSQAPEGDVISMPLHTSEELGFSKFVSA

**Important features:**

**Signal peptide:**

amino acids 1-17

**Transmembrane domain:**

amino acids 248-269

**N-glycosylation site:**

amino acids 96-99

**Fibrinogen beta and gamma chains C-terminal domain:**

amino acids 104-113

**Ig like V-type domain:**

amino acids 13-128

62/392

**FIGURE 61**

CGGGCCAGCCTGGGGCGGCCGGCCAGGAACCACCCGTTAAGGTGTCTTCTCTTTAGGGAT  
GGTGAGGTTGGAAAAAGACTCCTGTAAACCTCCTCCAGGATGAACCACCTGCCAGAAGAC  
ATGGAGAACGCTCTCACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCC  
ATCAACCCACACAACCTCATGGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGC  
ATATCTGATGTCAGGAGGACTTTCTGTTTGTTTGTACCTTTGACCTCTTATTCGTAACA  
TTACTGTGGATAATAGAGTTAAATGTGAATGGAGGCATTGAGAACACATTAGAGAAGGAG  
GTGATGCAGTATGACTACTATTCTTCATATTTTGATATATTTCTTCTGGCAGTTTTCGA  
TTTAAAGTGTTAATACTTGACATATGCTGTGTGCAGACTGCGCCATTGGTGGGCAATAGCG  
TTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAGTGATCCTTTTGAAGCTTTTC  
TCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTTCATTTCATCCTTGCCTGGATTGAG  
ACGTGGTTCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAACAGACTCCTG  
ATAGTTCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATGGTCAG  
TTTTATTCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT  
GAGAAACCACTTTTAGAACTATGAGTACTACTTTTGTTAAATGTGAAAAACCCTCACAGA  
AAGTCATCGAGGCAAAAAGAGGCAGGCAGTGGAGTCTCCCTGTGACAGTAAAGTTGAAA  
TGGTGACGTCCACTGCTGGCTTTATTGAACAGCTAATAAAGATTTATTTATTGTAATACC  
TCACAAACGTTGTACCATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAA  
TGTCATGATTCATCCTCTCTTCAGTGAGACTGAGCCTGATGTGTTAACAATAGGTGAAG  
AAAGTCTTGTGCTGTATTCCTAATCAAAAGACTTAATATATTGAAGTAACACTTTTTTAG  
TAAGCAAGATACCTTTTTTATTTCAATTCACAGAATGGAATTTTTTTTGTTCATGTCTCAG  
ATTTATTTTGTATTTCTTTTTTAACTCTACATTTCCCTTGTTTTTTAACTCATGCACA  
TGTGCTCTTTGTACAGTTTTTAAAAAGTGTAATAAAATCTGACATGTCAATGTGGCTAGTT  
TTATTTTTCTTGTTTTGCATTATGTGTATGGCCTGAAGTGTTGGACTTGCAAAAGGGGAA  
GAAAGGAATTGCGAATACATGTAAAATGTCACCAGACATTTGTATTATTTTTATCATGAA  
ATCATGTTTTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGCACAAA  
ATGACTTAAACCATTCATATCATGTTTCCTTTGCGTTCAGCCAATTTCAATTAAATGAA  
CTAAATTAAAAA

63/392

**FIGURE 62**

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVT  
FDLLFVTLWLWIELNVNGGIENTLEKEVMQYDYSSYFDIFLLAVFRFKVLILAYAVCRL  
RHWWAIALTTAVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLPQEA  
EEENRLLIVQDASERAALIPGGLSDGQFYSPPESEAGSEEAEKQDSEKPLLEL

**Important features of the protein:**

**Signal peptide:**

amino acids 1-20

**Transmembrane domains:**

amino acids 54-72, 100-118, 130-144, 146-166

**N-myristoylation sites:**

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

64/392

**FIGURE 63**

GCGCCGGGAGCCCATCTGCCCCCAGGGGACGCGGGCGCGGGGCCGGCTCCCGCCCCGGCAC  
ATGGCTGCAGCCACCTCGCGCGCACCCCCGAGGCGCGCGCCAGCTCGCCCCGAGGTCCGT  
CGGAGGCGCCCCGGCCGCCCCGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGGTCCG  
GGGATCGGGATGTCCTCTCTCTCTCTGCTAGTTTCTACTATGTTGGAACCTTG  
GGGACTCACACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACCTTGCCCTGCCACCAT  
CAACTGGGGCTTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAA  
GGGAACCAAAAAGTGGTGATCACTTACTCCAGTCGTCTACAATAACTTGACTGAG  
GAACAGAAGGGCCGAGTGGCCTTTGCTTCCAATTTCTGGCAGGAGATGCCTCCTTGACAG  
ATTGAACCTCTGAAGCCCACTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTCAGGG  
CGCTACGTGTGGAGCCATGTCATCTTAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGT  
GAGTTGGAAGGAGAGCTGACAGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCT  
GGCACAGAGCCCATTTGTGTATTACTGGCAGCGAATCCGAGAGAAAGAGGGAGAGGATGAA  
CGTCTGCCTCCCAAATCTAGGATTGACTACAACCACCCTGGACGAGTTCTGCTGCAGAAT  
CTTACCATGTCTACTCTGGACTGTACAGTGACAGCAGGCAACGAAGCTGGGAAGGAA  
AGCTGTGTGGTGCGAGTAACTGTACAGTATGTACAAAGCATCGGCATGGTTGCAGGAGCA  
GTGACAGGCATAGTGGCTGGAGCCCTGCTGATTTTCTCTTGGTGTGGCTGCTAATCCGA  
AGGAAAGACAAAGAAAGATATGAGGAAGAAGAGAGACCTAATGAAATTCGAGAAGATGCT  
GAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCTCCTCTTCTCAGGCTCTCGGAGCTCA  
CGCTCTGGTTCTTCTCCACTCGCTCCACAGCAAATAGTGCCTCACGCAGCCAGCGGACA  
CTGTCAACTGACGCAGCACCCAGCCAGGGCTGGCCACCCAGGCATACAGCCTAGTGGGG  
CCAGAGGTGAGAGGTCTTGAACCAAAGAAAGTCCACCATGCTAATCTGACCAAAGCAGAA  
ACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAAACGGTCTGAATTACAATG  
GACTTGACTCCCACGCTTTCTAGGAGTCAGGGTCTTTGGACTCTTCTCGTCATTGGAGC  
TCAAGTCACCAGCCACACAACCAGATGAGAGGTCACTAAGTAGCAGTGAGCATTGCACG  
GAACAGATTGAGATGAGCATTTTCTTTATACAATACCAAACAAGCAAAAGGATGTAAGCT  
GATTCATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGGAAAGCAGGAG  
TCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTGAGGTGAAT  
ATACCTAAAACTTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTACAATTTTCAAG  
AGGAAATGGGATGCTGTTTGTAAATTTTCTATGCATTTCTGCAAACTTATTGGATTATTA  
GTTATTCAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTACTG  
AGCTAACCACCTTCTAAGAACTCCAAAAAGGAAACATGTGTCTTCTATTTCTGACTTAAC  
TTCATTTGTCATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGA  
GAAGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACATTTTGTATTGAGCCCCAAA  
TAACTATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTCCATCTTCAT  
GATGTTATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCC  
CCTCAAATCAGATGCCTCTAAGGACTTTCTGCTAGATATTTCTGGAAGGAGAAAATACA  
ACATGTCATTTATCAACGTCCTTAGAAAGAATTTCTAGAGAAAAAGGGATCTAGGAAT  
GCTGAAAGATTACCCAACATACCATTATAGTCTCTTCTTTCTGAGAAAAATGTGAAACCAG  
AATTGCAAGACTGGGTGGACTAGAAAGGGAGATTAGATCAGTTTTCTCTTAATATGTCAA  
GGAAGGTAGCCGGGCATGGTGCCAGGCACCTGTAGGAAAAATCCAGCAGGTGGAGGTTGCA  
GTGAGCCGAGATTATGCCATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC



65/392

**FIGURE 64**

MSLLLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQ  
KVVITYSSRHVYNNLTEEQGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYV  
WSHVILKVLVRPSKPKCELEGELETEGSDTLQCESSSGTEPIVYYWQRIREKEGEDERLP  
PKSRIDYNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTQYVQSIGMVAGAVTG  
IVAGALLIFLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRS  
SSSTRSTANSASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTP  
SMIPSQSRAFQTV

**Signal sequence:**  
amino acids 1-16

**Transmembrane domain:**  
amino acids 232-251

66/392

**FIGURE 65**

GTCGGGGCTGCGCGACGGCGCAGGGGCTGCGGGGAGCGCCGCGCAGGCCGTGCAGTTCCT  
AGCGAGGAGGCGCCGCCGCCATTGCCGCTCTCTCGGTGAGCGCAGCCCCGCTCTCCGGGC  
CGGGCCTTCGCGGGCCACGGCGCCATGGGGCCAGTGCGGCATCACCTCCTCCAAGACCGT  
GCTGGTCTTTCTCAACCTCATCTTCTGGGGGGCAGCTGGCATTATGCTATGTGGGAGC  
CTATGTCTTCATCACTTATGATGACTATGACCACTTCTTTGAAGATGTGTACACGCTCAT  
CCCTGCTGTAGTGATCATAGCTGTAGGAGCCCTGCTTTTCATCATTGGGCTAATTGGCTG  
CTGTGCCACAATCCGGGAAAGTCGCTGTGGACTTGCCACGTTTGTATCATCCTGCTCTT  
GGTTTTTGTACAGAAGTTGTTGTAGTGGTTTTTGGGATATGTTTACAGAGCAAAGGTGGA  
AAATGAGGTTGATCGCAGCATTACAGAAAGTGTATAAGACCTACAATGGAACCAACCCTGA  
TGCTGCTAGCCGGGCTATTGATTATGTACAGAGACAGCTGCATTGTTGTGGAATTACAA  
CTACTCAGACTGGGAAAATACAGATTGGTTCAAAGAAACCAAAACCAGAGTGTCCCTCT  
TAGCTGCTGCAGAGAGACTGCCAGCAATTGTAATGGCAGCCTGGCCCCACCTTCCGACCT  
CTATGCTGAGGGGTGTGAGGCTCTAGTAGTGAAGAAGCTACAAGAAATCATGATGCATGT  
GATCTGGGCCGCACTGGCATTTCAGCTATTTCAGCTGCTGGGCATGCTGTGTGCTTGCAT  
CGTGTGTGTCAGAAGGAGTAGAGATCCTGCTTACGAGCTCCTCATCACTGGCGGAACCTA  
TGCATAGTTGACAACTCAAGCCTGAGCTTTTTTGGTCTTGTCTGATTGGAAGGTGAATT  
GAGCAGGTCTGCTGCTGTTGGCCTCTGGAGTTCATTTAGTTAAAGCACATGTACACTGGT  
GTTGGACAGAGCAGCTTGGCTTTTCATGTGCCACCTACTTACCTACTACCTGCGACTTT  
CTTTTTCCTTGTTCTAGCTGACTCTTCATGCCCCTAAGATTTTAAGTACGATGGTGAACG  
TTCTAATTTTCAAGAACCAATTGCGAGTCATGTAGTGTGGTAGAATTAAAGGAGGACACGAG  
CCTGCTTCTGTTACCTCCAAGTGGTAACAGGACTGATGCCGAAATGTCACCAGGTCCTTT  
CAGTCTTCACAGTGGAGAACTCTTGGCCAAAGTTTTTTCGGGGAGGAGGAGGAAACCAG  
CTTTCTGGTTAAGGTTAACACCAGATGGTGGCCCTCATTTGGTGTCTTTTAAAAAATATT  
TACTGTAGTCCAATAAGATAGCAGCTGTACAAAATGACTAAAATAGATTGTAGGATCATA  
TGGCGTATATCTTGGTTCATCTTCAAATCAGAGACTGAGCTTTGAACTAGTGGTTTTT  
AATCAAAGTTGGCTTTATAGGAGGAGTATAATGTATGCACTACTGTTTTTAAAGAATTAG  
TGTGAGTGTGTTTTTGTATGAATGAGCCCATTCATGGTAAGTCTTAAGCTTGTGGAAT  
AATGTACCCATGTAGACTAGCAAAATAGTATGTAGATGTGATCTCAGTTGTAAATAGAAA  
AATCTAATTCATAAACTCTGTATCAGCCCCCAAAAAAAAAAAAAAAAAA

67/392

**FIGURE 66**

MGQCGITSSKTVLVFLNLI FWGAAGILCYVGAYVFITYDDYDHHFFEDVYTLIPAVVIIAV  
GALLFIIGLIGCCATIRESRCGLATFVIIILLVFVTEV VVVVLGYVYRAKVENEVDRSIQ  
KVYKTYNGTNPDAASRAIDYVQRQLHCCGIHNYSDWENTDWFKETKNQSVPLSCCRETAS  
NCNGSLAHPSDLYAEGCEALVVKLQEIMMHVIWAALAFAAIQLLGMLCACIVLCRRSRD  
PAYELLITGGTYA

Signal peptide:  
none

Type II transmembrane domain:  
11-38

Other transmembrane domains:  
48-68, 87-107, 208-235

N-glycosylation site:  
127-131, 152-156, 167-171, 183-187

Tyrosine kinase phosphorylation site:  
236-244

N-myristoylation site:  
5-11, 68-74, 71-77, 226-232

Prokaryotic membrane lipoprotein lipid attachment site:  
62-73, 221-232

Transmembrane 4 family proteins:  
7-35, 56-106

68/392

**FIGURE 67**

GCGGCACCTGGAAGATGCGCCCATTTGGCTGGTGGCCTGCTCAAGGTGGTGTTCGTGGTCT  
TCGCCTCCTTGTGTGCCTGGTATTGCGGGTACCTGCTCGCAGAGCTCATTCCAGATGCAC  
CCCTGTCCAGTGCTGCCATATAGCATCCGCAGCATCGGGGAGAGGCCTGTCTCAAAGCTC  
CAGTCCCCAAAAGGCAAAAATGTGACCACTGGACTCCCTGCCCATCTGACACCTATGCCT  
ACAGGTTACTCAGCGGAGGTGGCAGAAGCAAGTACGCCAAAATCTGCTTTGAGGATAACC  
TACTTATGGGAGAACAGCTGGGAAATGTTGCCAGAGGAATAAACATTGCCATTGTCAACT  
ATGTAACCTGGGAATGTGACAGCAACACGATGTTTTGATATGTATGAAGGCGATAACTCTG  
GACCGATGACAAAGTTTATTCAGAGTGCTGCTCCAAAATCCCTGCTCTTCATGGTGACCT  
ATGACGACGGAAGCACAGACTGAATAACGATGCCAAGAATGCCATAGAAGCACTTGGAA  
GTAAAGAAATCAGGAACATGAAATTCAGGTCTAGCTGGGTATTTATTGCAGCAAAAGGCT  
TGGAACCTCCCTTCCGAAATTCAGAGAGAAAAGATCAACCACTCTGATGCTAAGAACAACA  
GATATTCCTGGCTGGCCTGCAGAGATCCAGATAGAAGGCTGCATACCCAAAGAACGAAGCT  
GACACTGCAGGGTCCTGAGTAAATGTGTTCTGTATAAACAAATGCAGCTGGAATCGCTCA  
AGAATCTTATTTTTCTAAATCCAACAGCCCATATTTGATGAGTATTTTGGGTTTGTGTA  
AACCAATGAACATTTGCTAGTTGTATCAAATCTTGGTACGCAGTATTTTTATACCAGTAT  
TTTATGTAGTGAAGATGTCAATTAGCAGGAACTAAAATGAATGGAAATTCCTAAAAAAA  
AAA

69/392

**FIGURE 68**

MRPLAGGLLKVVVFVVFASLCAWYSGYLLAELIPDAPLSSAAYSIRSIGERFVLKAPVPKR  
QKCDHWTPCPSDTYAYRLLSGGGRSKYAKICFEDNLLMGEQLGNVARGINIAIVNYVTGN  
VTATRCFDMYEGDNSGPMKFIQSAAPKSLLFMVTYDDGSTRLNNDAKNAIEALGSKEIR  
NMKFRSSWVFIAAKGLELPSEIQREKINHSDAKNNRYSGWPAEIQIEGCI PKERS

**Signal sequence:**

amino acids 1-20

**N-glycosylation sites:**

amino acids 120-124, 208-212

**Glycosaminoglycan attachment site:**

amino acids 80-84

**N-myristoylation sites:**

amino acids 81-87, 108-114, 119-125

70/392

**FIGURE 69**

ACACAACTTTACACCTGAATGAACGCCAAACCTCTATGGATATATAAAGGGAAGCTTGAG  
GAGGAATTTACAGTTACAGTGCAGAAGCAGAAGCAAAGAATTAACCAGCTCTTCAGTC  
AAGCAAATCCTCTACTCACCATGCTTCCTCCTGCCATTTCATTTCTATCTCCTTCCCCCTTG  
CATGCATCCTAATGAAAAGCTGTTTGGCTTTTAAAAATGATGCCACAGAAATCCTTTATT  
CACATGTGGTTAAACCTGTTCCAGCACACCCAGCAGCAACAGCACGTTGAATCAAGCCA  
GAAATGGAGGCAGGCATTTAGTAACACTGGACTGGATCGGAACACTCGGGTTCAAGTGG  
GTTGCCGGGAAGTGGCTTCCACCAAATACATCTCTGATGGCCAGTGCACCAGCATCAGCC  
CTCTGAAGGAGCTGGTGTGTGCTGGCGAGTGCTTGCCCCTGCCAGTGCTCCCTAACTGGA  
TTGGAGGAGGCTATGGAACAAAGTACTGGAGCAGGAGGAGCTCCCAGGAGTGGCGGTGTG  
TCAATGACAAAACCCGTACCCAGAGAATCCAGCTGCAGTGCCAAGATGGCAGCACACGCA  
CCTACAAAATCACAGTAGTCACTGCCTGCAAGTGCAAGAGGTACACCCGGCAGCACAAACG  
AGTCCAGTCACAACCTTTGAGAGCATGTACCTGCCAAGCCAGTCCAGCATCACAGAGAGC  
GGAAAAGAGCCAGCAAATCCAGCAAGCACAGCATGAGTTAGAACTCAGACTCCATAACT  
AGACTTACTAGTAACCATCTGCTTTACAGATTTGATTGCTTGGAAGACTCAAGCCTGCCA  
CTGCTGTTTTCTCACTTGAAAGTATATGCTTTCTGCTTTGATCAAACCCAGCAAGCTGTC  
TTAAGTATCAGGACCTTCTTTGGGAATAGTTTTTCCTTTTAAAGTTTTTCAAGATGTAGG  
TATATCCATGAATGCAATTTGCATTTAAATTCCACGTATCCCTGTAGTTTAAATTCCTCA  
TTGGTCTTAAAAGACTGTTGATACTATAAACATCAGTGAATCAATTATATTTTAAACA  
GAAAAGGGCTT

71/392

**FIGURE 70**

MLPPAIHFYLLPLACILMKSCLA FKNDAT EILYSHVVKPVP AHPSSNSTLNQARNGGRHF  
SNTGLDRNTRVQVGCRELRSTKYISDGQCT SISPLKELVCAGECLPLPVL PNWIGGGYGT  
KYWSRRSSQEWRCVNDKTRTQRIQLQCQDGSTR TYKITVVTACKCKRYTRQHNESSHNFE  
SMSPAKFPVQHHRRERKRASKSSKHSMS

**Signal sequence:**

1-23

**Transmembrane domain:**

None

**N-glycosylation site:**

47-50, 173-176

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

125-128, 166-169, 195-198

**N-myristoylation site:**

64-69, 87-92, 115-120, 116-121, 150-155

72/392

**FIGURE 71**

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAA  
AGGCTCCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTC  
CGAATCAGTAGGTGACCCCGCCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCCGA  
CCTCGTGCGGCCAAGACGTGGATGTTCCCTGCTCTTGCTGGGGGGAGCCTGGGCAGGACAC  
TCCAGGGCACAGGAGGACAAGGTGCTGGGGGGTCA TGAGTGCCAACCCCATTCGCAGCCT  
TGGCAGGCGGCCTTGTTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTG TAGGTGGC  
AACTGGGTCTTACAGCTGCCCACTGTAAAAAACCGAAATACACAGTACGCCTGGGAGAC  
CACAGCCTACAGAAATAAAGATGGCCCAGAGCAAGAAATACCTGTGGTTCAGTCCATCCCA  
CACCCCTGCTACAACAGCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAA  
CTGCGTGACCAGGCATCCCTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGC  
ACCCAGCCTGGCCAGAAGTGCACCGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAG  
AATTTTCCTGACACTCTCAACTGTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAG  
GATGCTTACCCGGGGCAGATCACAGATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCT  
GACACGTGCCAGGGCGATTCTGGAGGCCCCCTGGTGTGTGATGGTGCACTCCAGGGCATC  
ACATCCTGGGGCTCAGACCCCTGTGGGAGGTCCGACAAACCTGGCGTCTATACCAACATC  
TGCCGCTACCTGGACTGGATCAAGAAGATCATAGGCAGCAAGGGCTGGATTCTAGGATAAG  
CACTAGATCTCCCTTAATAAACTCACAACCTCTCTGGTTC



73/392

**FIGURE 72**

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGG  
VLVGGNWVLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIPHPCYNSSDVEDHNHD  
LMLLQLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFDTLNCAEVKIFP  
QKKCEDAYPGQITDGMVCAGSSKGADTCQGDSSGGLVCDGALQGITSWGSDDPCGRSDKPG  
VYTNICRYLDWIKKIIGSKG

**Important Features:****Signal peptide:**

amino acids 1-23

**Transmembrane domain:**

amino acids 51-71

**N-glycosylation site:**

amino acids 110-113

**Serine proteases, trypsin family, histidine active site:**

amino acids 69-74 and 207-217

**Tyrosine kinase phosphorylation site:**

amino acids 182-188

**Kringle domain proteins motif:**

amino acids 205-217

74/392

**FIGURE 73**

CTCGGGCGCGCACAGGCAGCTCGGTTTGGCCCTGCGATTGAGCTGCGGGTTCGCGGCCGGCG  
CCGGCCTCTCCAATGGCAAATGTGTGTGGCTGGAGGCGAGCGCGAGGCTTTCGGCAAAGG  
CAGTCGAGTGTTTGCAGACCGGGGCGAGTCCTGTGAAAGCAGATAAAAGAAAACATTTAT  
TAACGTGTCAATTACGAGGGGAGCGCCCGGCCGGGGCTGTCGCACTCCCCGCGGAACATTT  
GGCTCCCTCCAGCTCCGAGAGAGGAGAAGAAGAAAGCGGAAAAGAGGCAGATTACGTCG  
TTTCCAGCCAAGTGGACCTGATCGATGGCCCTCCTGAATTTATCACGATATTTGATTTAT  
TAGCGATGCCCCCTGGTTTGTGTGTTACGCACACACACGTGCACACAAGGCTCTGGCTCG  
CTTCCCTCCCTCGTTTCCAGCTCCTGGGCGAATCCACATCTGTTTCAACTCTCCGCCGA  
GGGCGAGCAGGAGCGAGAGTGTGTGCAATCTGCGAGTGAAGAGGGACGAGGGAAAAGAAA  
CAAAGCCACAGACGCAACTTGAGACTCCCGCATCCCAAAGAAGCACCAGATCAGCAAAA  
AAAGAAGATGGGCCCCCGAGCCTCGTGCTGTGCTTGCTGTCCGCAACTGTGTTCTCCCT  
GCTGGGTGGAAGCTCGGCCTTCTGTGCGACCACCGCTGAAAGGCAGGTTTCAGAGGGA  
CCGCAGGAACATCCGCCCAACATCATCTGGTGCTGACGGACGACCAGGATGTGGAGCT  
GGGTTCATGCAGGTGATGAACAAGACCCGGCGCATCATGGAGCAGGGCGGGGCGCACTT  
CATCAACGCCTTCGTGACCACACCCATGTGCTGCCCCTCACGCTCCTCCATCCTCACTGG  
CAAGTACGTCCACAACCACAACACCTACACCAACAATGAGAACTGCTCCTCGCCCTCCTG  
GCAGGCACAGCACGAGAGCCGCACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGAC  
AGCTTTCCTTCGGGAAGTATCTTAATGAATACAACGGCTCCTACGTGCCACCCGGCTGGAA  
GGAGTGGGTTCGACTCCTTAAAACTCCCGCTTTTATAACTACACGCTGTGTGCGAACGG  
GGTGAAGAGAAGCACGGCTCCGACTACTCCAAGGATTACCTCACAGACCTCATCACC  
TGACAGCGTGAGCTTCTTCCGCACGTCCAAGAAGATGTACCCGCACAGGCCAGTCCTCAT  
GGTCATCAGCCATGCAGCCCCCACGGCCCTGAGGATTCAGCCCCACAATATTCACGCCT  
CTTCCCAAACGCATCTCAGCACATCACGCCGAGCTACAACCTACGCGCCCAACCCGGACAA  
ACACTGGATCATGCGCTACACGGGGCCCATGAAGCCCATCCACATGGAATTCACCAACAT  
GCTCCAGCGGAAGCGCTTGACAGCCCTCATGTGCGGTGGACGACTCCATGGAGACGATTTA  
CAACATGCTGGTTGAGACGGGCGAGCTGGACAACACGTACATCGTATACACCGCCGACCA  
CGGTTACCACATCGGCCAGTTTGGCCTGGTGAAAGGGGAAATCCATGCCATATGAGTTTGA  
CATCAGGGTCCCGTTCTACGTGAGGGGGCCCCAACGTGGAAGCCGGCTGTCTGAATCCCCA  
CATCGTCTCAACATTGACCTGGCCCCCACCATCCTGGACATTGCAGGCCTGGACATACC  
TGCGGATATGGACGGGAAATCCATCCTCAAGCTGCTGGACACGGAGCGGCCGGTGAATCG  
GTTTCACTTGAAAAAGAAGATGAGGGTCTGGCGGGACTCCTTCTTGGTGGAGAGAGGCAA  
GCTGCTACACAAGAGAGACAATGACAAGGTGGACGCCAGGAGGAGAACTTTCTGCCCAA  
GTACCAGCGTGTGAAGGACCTGTGTGAGCGTGTGAGTACCAGACGGCGTGTGAGCAGCT  
GGGACAGAAGTGGCAGTGTGTGGAGGACGCCACGGGGAAGCTGAAGCTGCATAAGTGCAA  
GGGCCCCATGCGGCTGGGCGGCAGCAGAGCCCTCTCCAACCTCGTGCCCAAGTACTACGG  
GCAGGGCAGCGAGGCCTGCACCTGTGACAGCGGGGACTACAAGCTCAGCCTGGCCGACG  
CCGGAAAAAATCTTTCAAGAAGAAGTACAAGGCCAGCTATGTCCGCAGTCGCTCCATCCG  
CTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCACGTAGGCCTGGGTGATGCCGCCA  
GCCCCGAAACCTCACCAAGCGGCACTGGCCAGGGGCCCTGAGGACCAAGATGACAAGGA  
TGGTGGGGACTTCAGTGGCACTGGAGGCCTTCCCGACTACTCAGCCGCCAACCCCATTA  
AGTGACACATCGGTGCTACATCCTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTA  
CAAGTCCCTGCAGGCCTGGAAAGACCACAAGCTGCACATCGACCACGAGATTGAAACCT  
GCAGAACAAAATTAAGAACCTGAGGGAAGTCCGAGGTACCTGAAGAAAAAGCGGCCAGA  
AGAATGTGACTGTCAAAAATCAGCTACCACACCCAGCACAAAGGCCGCCTCAAGCACAG  
AGGCTCCAGTCTGCATCCTTTCAGGAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTT  
GCGGGAGCAGAAGCGCAAGAAGAACTCCGCAAGCTGCTCAAGCGCCTGCAGAACACGA  
CACGTGCAGCATGCCAGGCCTCACGTGCTTACCCACGACAACCAGCACTGGCAGACGGC

75/392

GCCTTTCTGGACACTGGGGCCTTTCTGTGCCTGCACCAGCGCCAACAATAACACGTACTG  
GTGCATGAGGACCATCAATGAGACTCACAATTCCTCTTCTGTGAATTTGCAACTGGCTT  
CCTAGAGTACTTTTGATCTCAACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACT  
GGACAGGGATGTCCTCAACCAGCTACACGTACAGCTCATGGAGCTGAGGAGCTGCAAGGG  
TTACAAGCAGTGTAACCCCCGGAAGTTCGAAACATGGACCTGGATGGAGGAAGCTATGAGCA  
ATACAGGCAGTTTCAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACT  
GGGACAACTGTGGGAAGGCTGGGAAGGTTAAGAAACAACAGAGGTGGACCTCCAAAAACA  
TAGAGGCATCACCTGACTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACC  
TGTGCTATTGGCCAGGAGGCTGAGAAAGCAAGCACGCACTCTCAGTCAACATGACAGAT  
TCTGGAGGATAACCAGCAGGAGCAGAGATAACTTCAGGAAGTCCATTTTGGCCCTGCTT  
TTGCTTTGGATTATACCTCACCAGCTGCACAAAATGCATTTTTTCGTATCAAAAAGTCAC  
CACTAACCCCTCCCCCAGAAGCTCACAAAGGAAAACGGAGAGAGCGAGCGAGAGAGATTTT  
CTTGGAATTTCTCCCAAGGGCGAAAGTCATTGGAATTTTAAATCATAGGGGAAAAGCA  
GTCCTGTTCTAAATCCTCTTATTCTTTTGGTTTGTCAAAAGAAGGAACATAAGAAGCAGG  
ACAGAGGCAACGTGGAGAGGCTGAAAACAGTGCAGAGACGTTTGACAATGAGTCAGTAGC  
ACAAAAGAGATGACATTTACCTAGCACTATAAACCCCTGGTTGCCTCTGAAGAACTGCCT  
TCATTGTATATATGTGACTATTTACATGTAATCAACATGGGAACCTTTAGGGGAACCTAA  
TAAGAAATCCCAATTTTCAGGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAA  
GAAAAA

76/392

**FIGURE 74**

MGPPSLVLCLLSATVFSLGGSSAFLSHRLKGRFQDRRNIRPNIILVLTDDQDVELGS  
MQVMNKTRRIMEQGGAHFINAFVTPMCCPSRSSILTGKYVHNHNTYTNNENCSSPSWQA  
QHESRTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVK  
EKHGSDYSKDYLTDLITNDSVSFFRTSKKMPHPRVLMVISHAAPHGPEDSAPQYSRLFP  
NASQHITPSYNYAPNPDKHWMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNM  
LVETGELDNTYIVYTADHGYHIGQFGLVKGKSMPIEFDIRVPFYVRGPNVEAGCLNPHIV  
LNIDLAPTILDIAGLDIPADMKGKSLKLLDTERPVMRFLKKKMRVWRDSFLVERGKLL  
HKRDNDKVDAQEENFLPKYQRVKDLQRAEYQTACEQLGQKWQCVEDATGKLKLHKCKGP  
MRLGGSRALSNLVPKYYGQSEACTCDSDYKLSLAGRRKKLFKKKYKASYVRSRSIRSV  
AIEVDGRVYHVLGDAQAQRNLTKRHWPGAPEDQDDKDGGDFSGTGGLPDYSAANPIKVT  
HRCYILENDTVQCDLDLYKSLQAWKDHLHIDHEIETLQNKIKNLREVRGHLKKKRPEEC  
DCHKISYHTQHKGRCLKHRGSSLHPFRKGLQEKDKVWLLREQKRKKKLRKLLKRLQNNDTC  
SMPGLTCFTHDNQHWQTAPFWTLGPFCACTSANNNNTYWCMTINETHNLFCEFATGFLE  
YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMDLDGGSYEQYR  
QFQRRKWPEMKRPSSKSLGQLWEGWEG

**Important features:****Signal peptide:**

amino acids 1-17

**Sulfatases signature 1:**

amino acids 86-99

**Homologous region to sulfatase:**

amino acids 87-106, 133-146, 216-229, 291-320, 365-375

**N-glycosylation sites:**amino acids 65-69, 112-116, 132-136, 149-153, 171-175,  
198-202, 241-245, 561-565, 608-612, 717-721, 754-758,  
764-768

77/392

**FIGURE 75**

CCCACGCGTCCGCCCACGCGTCCGGTGGACTATGGGCCAGTTTTTGTGCAAGAACCAGAT  
GATATTATTTTTTCCAAGTCTGATGAAAAGAAGGTAGCATTGAATTGTGAAGTTCGT  
GGCAATCCAGTTCCAGTTACAGATGGCTTCGAAATGGAACAGAAATAGATCTGGAAAGT  
GATTATCGCTACAGTTTGATAGATGGCACCTTCATTATAAGCAATCCAAGTGAAGCAAAG  
GATTCTGGTCATTATCAGTGTTTAGCAACCAACACTGTGGGGAGTATTCTTAGTAGAGAA  
GCTACACTGCAGTTTGCTATCTGGGAAATTTTAGTGGCCGGACAAGAAGTGCAGTCTCT  
GTGAGGGAAGGCCAGGGTGTCTGTTCTGATGTGCTCTCCTCCGCCACATTCACCAGAGATC  
ATCTATAGCTGGGTATTTAATGAGTTCCTTCCTTTGTGGCGGAAGACAGCCGGCCGGTTC  
ATCTCCAGGAGACAGGCAACCTTTATATTTCTAAAGTCCAAACATCAGATGTTGGCAGC  
TATATTTGTCTGGTGAAAAACACAGTGACGAATGCTAGAGTCCTTAGTCTCCAACGCCA  
CTCACTCTGCGTAATGATGGTGTGATGGGAGAATATGAGCCGAAAATTGAGGTCCATTTT  
CCTTTCACGGTTACAGCTGCTAAAGGAACAACCTGTTAAGATGGAATGCTTTGCACTTGGC  
AACCCCGTTCCAACAATCACATGGATGAAGGTTAATGGTTATATTCCTAGTAAGGCACGT  
CTGCGGAAATCTCAGGCGGTGCTGGAAATACCGAATGTACAGCTGGATGATGCAGGCATT  
TATGAGTGCAGAGCTGAAAACCTCACGTGGAAAAAATTCCTTTCGTGGACAATTACAAGTA  
TACACCTACCCACACTGGGTAGAAAACTGAATGATACTCAGTTAGACAGTGGGAGCCCT  
CTCCGATGGGAATGTAAGGCTACTGGAAAACCCAGACCCACGTATCGTTGGCTGAAGAAT  
GGAGTACCCCTCTCACCTCAGAGTAGGGTTGAGATGGTTAATGGAGTATTGATGATCCAC  
AATGTGAATCAATCAGATGCTGGAATGTATCAGTGTTTGGCTGAAAATAAGTATGGAGCC  
ATTTACGCTAGTGCTGAGCTGAAGATTCTAGCTTCAGCTCCCACTTTTGCACTGAATCAA  
CTGAAGAAAACAATAATTGTTACCAAAGACCAAGAAGTTGTCATAGAGTGCAAACCCCAA  
GGCTCTCCAAACCAACCATCTCTTGGAAGAAAGGAGACAGAGCAGTTAGAGAAAACAAA  
AGAATAGCTATTCTTCCAGACGGGAGTCTACGGATCCTAAATGCTTCCAATCAGACGAG  
GGAAAGTACGTTTGCCGAGGGGAAAACGTCTTTGGTTCTGCTGAAAT

78/392

**FIGURE 76**

MCSPPPHSPEIIYSWVFNEFP SFVAEDSRRFISQETGNLYISKVQTS DVGSIYICLVKNTV  
TNARVLSPPPTPLTLRNDGVMGEYEPKIEVHF PFTVTAAGTTVKMECFALGNPVPTITWM  
KVNGYIPSKARLRKSQAVLEIPNVQLDDAGIYECRAENSRGKNSFRGQLQVYTYPHWVEK  
LNDTQLDSGSPLRWECKATGKPRPTYRWLKNGVPLSPQSRVEMVNGVLMIHNVNQSDAGM  
YQCLAENKYGAIYASAELKILASAPTFALNQLKKTIIIVTKDQEVVIECKPQGS PKPTISW  
KKGDRAVRE NKRIAILPDGSLRILNASKSDEGKYVCRGENVFGSAE

**Signal sequence:**

None

**Transmembrane domain:**

None

**N-glycosylation site:**

182-185, 234-237, 325-328

**Tyrosine kinase phosphorylation site:**

328-334

**N-myristoylation site:**

50-55, 150-155, 239-244, 250-255

**Immunoglobulin domain:**

2-56, 100-156, 189-245, 281-338

79/392

## FIGURE 77

GCTCCCAGCCAAGAACCTCGGGGCCGCTGCGCGGTGGGGAGGAGTTCCCCGAAACCCGGC  
CGCTAAGCGAGGCCTCCTCCTCCCGCAGATCCGAACGGCCTGGGCGGGGTACCCCGGGT  
GGGACAAGAAGCCCGCCGCTGCCTGCCCGGGCCCGGGAGGGGGCTGGGGCTGGGGCCGG  
AGGCGGGGTGTGAGTGGGTGTGTGCGGGGGCGGAGGCTTGATGCAATCCCGATAAGAAA  
TGCTCGGGTGTCTTGGGCACCTACCCGTGGGGCCCCGTAAGGCGCTACTATATAAGGCTGC  
CGGCCCCGAGCCCGCCGCGCGTCAGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCCACGA  
CCATCCCAACCCGGCACTCACAGCCCCGAGCGCATCCCGGTGCGCGCCAGCCTCCCCG  
ACCCCATCGCCGGAGCTGCGCCGAGAGCCCCAGGGAGGTGCCATGCGGAGCGGGTGTGT  
GGTGGTCCACGTATGGATCCTGGCCGGCCTCTGGCTGGCCGTGGCCGGGCGCCCCCTCGC  
CTTCTCGGACGCGGGGCCCCACGTGCATACGGCTGGGGCGACCCCATCCGCCTGCGGCA  
CCTGTACACCTCCGGCCCCCACGGGCTCTCCAGCTGCTTCTGCGCATCCGTGCGGACGG  
CGTTCGTGGACTGCGCGCGGGGCCAGAGCGCGCACAGTTTGCTGGAGATCAAGGCAGTCGC  
TCTGCGGACCGTGGCCATCAAGGGCGTGACAGCGTGCGGTACCTCTGCATGGGCGCCGA  
CGGCAAGATGCAGGGGCTGCTTCAGTACTCGGAGGAAGACTGTGCTTTCGAGGAGGAGAT  
CCGCCCAGATGGCTACAATGTGTACCGATCCGAGAAGCACCGCCTCCCGGTCTCCCTGAG  
CAGTGCCAAACAGCGGCAGCTGTACAAGAACAGAGGCTTTCTTCCACTCTCTCATTTCT  
GCCCATGCTGCCATGGTCCCAGAGGAGCCTGAGGACCTCAGGGGCCACTTGGAATCTGA  
CATGTTCTCTTCGCCCCCTGGAGACCGACAGCATGGACCCATTTGGGCTTGTACCCGACT  
GGAGGCCGTGAGGAGTCCCAGCTTTGAGAAGTAACTGAGACCATGCCCGGCCTCTTCAC  
TGCTGCCAGGGGCTGTGCTACCTGCAGCGTGGGGGACGTGCTTCTACAAGAACAGTCCTG  
AGTCCACGTTCTGTTTAGCTTTAGGAAGAAACATCTAGAAGTTGTACATATTCAGAGTTT  
TCCATTGGCAGTGCCAGTTTCTAGCCAATAGACTTGTCTGATCATAACATTGTAAGCCTG  
TAGCTTGCCCAGCTGCTGCTGGGCCCCCATTTCTGCTCCCTCGAGGTTGCTGGACAAGCT  
GCTGCACTGTCTCAGTTCTGCTTGAATACCTCCATCGATGGGGAACCTCACTTCCTTTGGA  
AAAATTCTTATGTCAAGCTGAAATTCTCTAATTTTTTCTCATCACTTCCCCAGGAGCAGC  
CAGAAGACAGGCAGTAGTTTTAATTTTCAAGAACAGGTGATCCACTCTGTAAAACAGCAGG  
TAAATTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGGACCATTG  
CCCTTCCCAAATCCCTCCAGGCCAGAACTGACTGGAGCAGGCATGGCCCACCAGGCTTCA  
GGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACTTGAGAATTTCCCTGTA  
GGCCAGTTCTGTGATGCTGTCTGAGAATAACTTGCTGTCCCGGTGTACCTGCTT  
CCATCTCCAGCCCCACCAGCCCTCTGCCCACCTCACATGCCTCCCATGGATTGGGGCCT  
CCCAGGCCCCCACCTTATGTCAACCTGCCTTCTTGTTCAAAAATCAGGAAAAGAAAAG  
ATTTGAAGACCCCAAGTCTTGTCAATAACTTGCTGTGTGGAAGCAGCGGGGAAGACCTA  
GAACCCTTTCCCCAGCACTTGGTTTTTCAACATGATATTTATGAGTAATTTATTTTGATA  
TGTACATCTCTTATTTTCTTACATTATTTATGCCCCAAATTATTTATGTATGTAAGT  
GAGGTTTGTTTTGTATATTAAATGGAGTTTGTTTGT

80/392

**FIGURE 78**

MRSGCVVHVWILAGLWLAVAGRPLAFSDAGPHVHYGWGDP IRLRHLYTSGPHGLSSCFL  
RIRADGVVDCARGQSAHSLL EIKAV ALRTVAIKGVHSVRYLCMGADGKMQLLQYSEEDC  
AFEEEIRPDGYNVYRSEK HRLPVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPEEPEDLR  
GHLESDMFSSPLETDSMDPFGLVTGLEAVRSPSF EK

**Signal peptide:**  
amino acids 1-22

**Casein kinase II phosphorylation site:**  
amino acids 78-82, 116-120, 190-194, 204-208

**N-myristoylation site:**  
amino acids 15-21, 54-60, 66-72, 201-207

**Prokaryotic membrane lipoprotein lipid attachment site:**  
amino acids 48-59



81/392

## FIGURE 79

CGGACGCGTGGGCGGACGCGTGGGCCTGGGCAAGGGCCGGGGCGCCGGGCGGAGCCACCTCTTCCC  
CTCCCCCGCTTCCCTGTGCGGCTCCGCTGGCTGGACGCGCTGGAGGAGTGGAGCAGCACCCGGCCG  
GCCCTGGGGGCTGACAGTCCGCAAAAGTTTGGCCCCGAAGAGGAAGTGGTCTCAAACCCCGGCAGGTG  
GCGACCAGGCCAGACCAGGGGCGCTCGCTGCCCTGCGGGCGGGCTGTAGGCGAGGGCGCGCCCCAGT  
GCCGAGACCCGGGGCTTCAAGAGCCGGCCCCGGGAGAGAAGAGTGC GGCGGGACGGAGAAAAACA  
ACTCCAAAGTTGGCGAAAGGCACCGCCCCCTACTCCCGGGCTGCCGCGCCCTCCCCGCCCCCAGCCC  
TGGCATCCAGAGTACGGGTGAGCCCCGGGCCATGGAGCCCCCTGGGGAGGCGGCACCAGGGAGCC  
TGGGCGCCCCGGGGCTCCGCGCGGACCCCATCGGGTAGACCACAGAAGCTCCGGGACCCCTTCCGGCA  
CCTCTGGACAGCCCAGGATGCTGTTGGCCACCTCCTCCTCCTCCTCCTTGGAGGCGCTCTGGCCC  
ATCCAGACCCGATTATTTTCCAAATCATGCTTGTGAGGACCCCCCAGCAGTGTCTTAGAAGTGC  
AGGGCACCTTACAGAGGCCCTGGTCCGGGACAGCCGCACCTCCCCTGCCAACTGCACCTGGCTCA  
TCCTGGGCGAGCAAGGAACAGACTGTCAACATCAGGTTCCAGAAGCTACACCTGGCCTGTGGCTCAG  
AGCGCTTAACCTACGCTCCCCCTCCAGCCACTGATCTCCCTGTGTGAGGCACCTCCCAGCCCTC  
TGCAGCTGCCCGGGGGCAACGTCAACATCACTTACAGCTATGCTGGGGCCAGAGCACCCATGGGGCC  
AGGGCTTCTGTCTCTCTACAGCCAAGATTGGCTGATGTGCCTGCAGGAAGAGTTTCAGTGCCCTGA  
ACCACCGCTGTGTATCTGTGTCCAGCGCTGTGATGGGGTTGATGCCCTGTGGCGATGGCTCTGATG  
AAGCAGGTTGCAGCTCAGACCCCTTCCCTGGCTGACCCCAAGACCCGTCCCCCTCCCTGCCCTTGCA  
ATGTCACCTTGGAGGACTTCTATGGGGTCTTCTCCTCCTCCTGGATATACACACCTAGCCTCAGTCT  
CCCACCCCCAGTCCCTGCCATTGGCTGTGGACCCCCATGATGGCCGGCGGTGGCCGTGCGCTTCA  
CAGCCCTGGACTTGGGCTTTGGAGATGCAGTGCATGTGTATGACGGCCCTGGGGCCCCCTGAGAGCT  
CCCGACTACTGCGTAGTCTCACCCACTTCAGCAATGGCAAGGCTGTCACTGTGGAGACACTGTCTG  
GCCAGGCTGTGTGCTTACCAACACAGTTGCTTGGAGCAATGGTCTGTGGCTTCAATGCCACCTACC  
ATGTGCGGGGCTATTGCTTGCCTTGGGACAGACCTGTGGCTTAGGCTTGGCTTGGGAGCTGGCG  
AAGGCCTAGGTGAGCGCTGCTACAGTGAGGCACAGCGCTGTGACGGCTCATGGGACTGTGCTGACG  
GCACAGATGAGGAGGACTGCCCAGGCTGCCACCTGGACACTTCCCCCTGTGGGGCTGCTGGCACCT  
CTGGTGCCACAGCCTGCTACCTGCCTGCTGACCGCTGCAACTACCAGACTTCTGTGCTGATGGAG  
CAGATGAGAGACGCTGTCGGCATTGCCAGCCTGGCAATTTCCGATGCCGGGACGAGAAGTGCGTGT  
ATGAGACGTGGGTGTGCGATGGGCAGCCAGACTGTGCGGACGGCAGTGATGAGTGGGACTGCTCCT  
ATGTTCTGCCCCGCAAGGTCATTACAGCTGCAGTCATTGGCAGCCTAGTGTGCGGCCTGCTCCTGG  
TCATCGCCCTGGGCTGCACCTGCAAGCTCTATGCCATTGCAACCCAGGAGTACAGCATCTTTGCCC  
CCCTCTCCCGGATGGAGGCTGAGATTGTGACAGCAGGCAGGCCACCCCTTCTACCGGCGAGCTCATG  
CCCAGGGTGCCATCCACCTGTAGAAGACTTTCCTACAGAGAATCCTAATGATAACTCAGTGTGG  
GCAACCTGCGTTCTCTGTACAGATCTTACGCCAGGATATGACTCCAGGAGGTGGCCAGGTGCC  
GCCGTGCTCAGCGGGGCGCTTGTGTCGACGCTGGTACGCGCTCTCCGCGCTGGGGCTTGCTCC  
CTCGAACCAACACCCCGGCTCGGGCCTCTGAGGCCAGATCCAGGTCACACCTTCTGCTGCTCCCC  
TTGAGGCCCTAGATGGTGGCACAGGTCCAGCCCGTGAGGGCGGGGCGAGTGGGTGGGCAAGATGGGG  
AGCAGGCACCCCCACTGCCCATCAAGGCTCCCCCTCCCATCTGCTAGCACGTCTCCAGCCCCCACTA  
CTGTCCCTGAAGCCCCAGGGCCACTGCCCTCACTGCCCTTAGAGCCATCACTATTGTCTGGAGTGG  
TGCAGGCCCTGCGAGGCCGCTGTTGCCAGCCTGGGGCCCCCAGGACCAACCCGAGCCCCCTG  
GACCCCAACAGCAGTCTTGGCCCTGGAAGATGAGGACGATGTGCTACTGGTGCCACTGGCTGAGC  
CGGGGGTGTGGGTAGCTGAGGCAGAGGATGAGCCACTGCTTACCTGAGGGGACCTGGGGGCTCTAC  
TGAGGCCTCTCCCCTGGGGGCTCTACTCATAGTGGCAACCTTTTAGAGGTGGGTGAGCCTCCCC  
TCCACCACTTCCCTTCCCTGTCCCTGGATTTTCAAGGACTTGGTGGGCTCCCGTTGACCTATGTAG  
CTGCTATAAAGTTAAGTGTCCCTCAGGCAGGGAGAGGGCTCACAGAGTCTCCTCTGTACGTGGCCA  
TGGCCAGACACCCAGTCCCTTACCACCACCTGCTCCCCACGCCACCACCTTTGGGTGGCTGTT  
TTTAAAAAGTAAAGTTCTTAGAGGATCATAGTCTGGACACTCCATCCTTGCCAAACCTCTACCCA  
AAAGTGGCCTTAAGCACCGGAATGCCAATTAAGTAGAGCCCTCCAGCCCCAAGGGGAGGATTTG  
GGCAGAACCTGAGGTTTTGCCATCCACAATCCCTCCTACAGGCGCTGGCTCACAAAAGAGTGCAA  
CAAATGCTTCTATTCCATAGCTACGGCATTGCTCAGTAAGTTGAGGTCAAAAATAAAGGAATCATA  
CATCTC

82/392

**FIGURE 80**

MLLATLLLLLLGGALAHPDRIIFPNHACEDPPAVLLEVQGTLQRPLVRDSRTSPANCTWL  
ILGSKEQTVTIRFQKLHLACGSERLTLRSPLOPLISLCEAPPSPLQLPGGNVTITYSYAG  
ARAPMGQGFLLSYSQDWLMCLQEEFQCLNHRCVSAVQRCDGVDACGDGSDEAGCSSDPFP  
GLTPRPVPSLPCNVTLEDFYGVFSSPGYTHLASVSHPQSCHWLLDPHDGRRRLAVRFTALD  
LGFGDAVHVYDGPPESSRLLRSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNGRGFNA  
TYHVRGYCLPWDRPCGLGSGLGAGEGLGERCYSEAQRCDGSWDCADGTDEEDCPGCPPGH  
FPCGAAGTSGATACYLPADRCNYQTFCADGADERRCRHCQPGNFRRCRDEKCVYETWVCDG  
QPD CADGSDEWDCSYVLPRKVITA AVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAPL  
SRMEAEIVQQAPPSSYGQLIAQGAIPPVEDFPTENPNDNSVLGNLRSLLQILRQDMTPGG  
GPGARRRQRGRMLMRRLLVRRLLRWGLLPRTNTPARASEARSQVTPSAAPLEALDGGTGPAR  
EGGAVGGQDGEQAPPLPIKAPLPSASTSPAPTTVPEAPGPLPSLPLEPSLLSGVVQALRG  
RLLPSLGP PGPTRSPPGPHTAVLALEDEDDVLLVPLAEPGVWVAEAEDEPLLT

**Important features:**

**Signal peptide:**

amino acids 1-16

**Transmembrane domain:**

amino acids 442-462

**LDL-receptor class A (LDLRA) domain proteins:**

amino acids 411-431, 152-171, 331-350 and 374-393

83/392

## FIGURE 81

CTTCTGTGCTGTTCCCTTCTTGCCTCTAACTTGTAACAAGACGTACTAGGACGATGCTAA  
TGAAAGTCAAAACCGCTGGGTTTTTGAAGGATCCTTGGGACCTCATGCACATTTGTG  
GAACTGGATGGAGAGATTTGGGGAAGCATGGACTCTTTAGCCAGCTTAGTTCTCTGTGG  
AGTCAGCTTGCTCCTTTCTGGAAGTGTGGAAGGTGCCATGGACTTGATCTTGATCAATTC  
CCTACCTCTTGTATCTGATGCTGAAACATCTCTCACCTGCATTGCCTCTGGGTGGCGCCC  
CCATGAGCCCATCACCATAGGAAGGGACTTTGAAGCCTTAATGAACCAGCACCAGGATCC  
GCTGGAAGTTACTCAAGATGTGACCAGAGAATGGGCTAAAAAGTTGTTTTGGAAGAGAGA  
AAAGGCTAGTAAGATCAATGGTGCTTATTTCTGTGAAGGGCGAGTTCGAGGAGAGGCAAT  
CAGGATACGAACCATGAAGATGCGTCAACAAGCTTCCTTCCCTACCAGCTACTTTAACTAT  
GACTGTGGACAAGGGAGATAACGTGAACATATCTTTCAAAAAGGTATTGATTAAAGAAGA  
AGATGCAGTGATTTACAAAAATGGTTCCTTCATCCATTAGTGCCCCGGCATGAAGTACC  
TGATATTCTAGAAGTACACCTGCCTCATGCTCAGCCCCAGGATGCTGGAGTGTACTCGGC  
CAGGTATATAGGAGGAAACCTCTTCACCTCGGCCTTCACCAGGCTGATAGTCCGGAGATG  
TGAAGCCCAGAAGTGGGGACCTGAATGCAACCATCTCTGTACTGCTTGTATGAACAATGG  
TGTCTGCCATGAAGATACTGGAGAATGCATTTGCCCTCCTGGGTTTATGGGAAGGACGTG  
TGAGAAGGCTTGTGAAGTGCACACGTTTTGGCAGAAGTTGTAAAGAAAGGTGCAGTGGACA  
AGAGGGATGCAAGTCTTATGTGTTCTGTCTCCCTGACCCCTATGGGTGTTCTGTGCCAC  
AGGCTGGAAGGGTCTGCAGTGCAATGAAGCATGCCACCCTGGTTTTTACGGGGCCAGATTG  
TAAGCTTAGGTGCAGCTGCAACAATGGGGAGATGTGTGATCGCTTCCAAGGATGTCTCTG  
CTCTCCAGGATGGCAGGGGCTCCAGTGTGAGAGAGAAGGCATACCGAGGATGACCCCAA  
GATAGTGGATTTGCCAGATCATATAGAAGTAAACAGTGGTAAATTTAATCCCATTTGCAA  
AGCTTCTGGCTGGCCGCTACCTACTAATGAAGAAATGACCCTGGTGAAGCCGGATGGGAC  
AGTGCTCCATCCAAAAGACTTTAACCATACGGATCATTTCTCAGTAGCCATATTACCAT  
CCACCGGATCCTCCCCCTGACTCAGGAGTTTGGGTCTGCAGTGTGAACACAGTGGCTGG  
GATGGTGGAAAAGCCCTTCAACATTTCTGTTAAAGTTCTTCCAAAGCCCTGAATGCCCC  
AAACGTGATTGACACTGGACATAACTTTGCTGTCAACATCAGCTCTGAGCCTTACTT  
TGGGGATGGACCAATCAAATCCAAGAAGCTTCTATACAAACCCGTTAATCACTATGAGGC  
TTGGCAACATATTTCAAGTGACAAATGAGATTGTTACACTCAACTATTTGGAACCTCGGAC  
AGAATATGAACTCTGTGTGCAACTGGTCCGTGCTGGAGAGGGTGGGGAAGGGCATCCTGG  
ACCTGTGAGACGCTTCACAAAGCTTCTATCGGACTCCCTCCTCCAAGAGGTCTAAATCT  
CCTGCCTAAAAGTCAGACCACTCTAAATTTGACCTGGCAACCAATATTTCCAAGCTCGGA  
AGATGACTTTTATGTTGAAGTGGAGAGAAGGTCTGTGCAAAAAGTGATCAGCAGAATAT  
TAAAGTTCCAGGCAACTTGACTTCGGTGCTACTTAACAACTTACATCCCAGGGAGCAGTA  
CGTGGTCCGAGCTAGAGTCAACACCAAGGCCAGGGGAATGGAGTGAAGATCTCACTGC  
TTGGACCCCTTAGTGACATTCTTCTCCTCAACCAGAAAACATCAAGATTTCCAACATTAC  
ACACTCCTCGGCTGTGATTTCTTGGACAATATTGGATGGCTATTCTATTCTTCTATTAC  
TATCCGTTACAAGGTTCAAGGCAAGAATGAAGACCAGCACGTTGATGTGAAGATAAAGAA  
TGCCACCATCATTCAGTATCAGCTCAAGGGCCTAGAGCCTGAAACAGCATACCAGGTGGA  
CATTTTTGTCAGAGAAACATAGGGTCAAGCAACCCAGCCTTTTCTCATGAACTGGTGAC  
CCTCCAGAACTCTCAAGCACCAGCGGACCTCGGAGGGGGGAAGATGCTGCTTATAGCCAT  
CCTTGGCTCTGCTGGAATGACCTGCCTGACTGTGCTGTTGGCCTTTCTGATCATATTGCA  
ATTGAAGAGGGCAAATGTGCAAAGGAGAATGGCCCAAGCCTTCCAAAACGTGAGGGAAGA  
ACCAGCTGTGCAGTTCAACTCAGGGACTCTGGCCCTAAACAGGAAGGTCAAAAACAACCC  
AGATCCTACAATTTATCCAGTGCTTGACTGGAATGACATCAAATTTCAAGATGTGATTGG  
GGAGGGCAATTTTGGCCAAGTTCTTAAGGCGCGCATCAAGAAGGATGGGTTACGGATGGA  
TGCTGCCATCAAAAGAATGAAAGAATATGCCTCCAAAGATGATCACAGGGACTTTGCAGG  
AGAAGTGAAGTTCTTTGTAAACTTGGACACCATCCAAACATCATCAATCTCTTAGGAGC

84/392

ATGTGAACATCGAGGCTACTTGTACCTGGCCATTGAGTACGCGCCCCATGGAAACCTTCT  
GGACTTCCTTCGCAAGAGCCGTGTGCTGGAGACGGACCCAGCATTTGCCATTGCCAATAG  
CACCGCGTCCACACTGTCTCCAGCAGCTCCTTCACTTCGCTGCCGACGTGGCCCGGGG  
CATGGACTACTTGAGCCAAAAACAGTTTATCCACAGGGATCTGGCTGCCAGAAACATTTT  
AGTTGGTGAAAACATATGTGGCAAAAATAGCAGATTTTGGATTGTCCCGAGGTCAAGAGGT  
GTACGTGAAAAAGACAATGGGAAGGCTCCCAGTGCGCTGGATGGCCATCGAGTCACTGAA  
TTACAGTGTGTACACAACCAACAGTGATGTATGGTCCTATGGTGTGTTACTATGGGAGAT  
TGTTAGCTTAGGAGGCACACCCCTACTGCGGGATGACTTGTGCAGAACTCTACGAGAAGCT  
GCCCCAGGGCTACAGACTGGAGAAGCCCCCTGAACTGTGATGATGAGGTGTATGATCTAAT  
GAGACAATGCTGGCGGGAGAAAGCCTTATGAGAGGCCATCATTTGCCAGATATTGGTGTC  
CTTAAACAGAATGTTAGAGGAGCGAAAGACCTACGTGAATACCACGCTTTATGAGAAGTT  
TACTTATGCAGGAATTGACTGTTCTGCTGAAGAAGCGGCCCTAGGACAGAACATCTGTATA  
CCCTCTGTTTCCCTTTCACTGGCATGGGAGACCCCTTGACAACTGCTGAGAAAACATGCCT  
CTGCCAAAGGATGTGATATATAAGTGTACATATGTGCTGGAATTCTAACAAGTCATAGGT  
TAATATTTAAGACACTGAAAAATCTAAGTGATATAAATCAGATTCTTCTCTCTCATTTTA  
TCCCTCACCTGTAGCATGCCAGTCCCGTTTCATTTAGTCATGTGACCACTCTGTCTTG  
TTTCCACAGCCTGCAAGTTCAGTCCAGGATGCTAACATCTAAAAATAGACTTAAATCTCA  
TTGCTTACAAGCCTAAGAATCTTTAGAGAAGTATACATAAGTTTAGGATAAAATAATGGG  
ATTTTCTTTTCTTTTCTCTGGTAATATTGACTTGTATATTTTAAGAAATAACAGAAAGCC  
TGGGTGACATTTGGGAGACATGTGACATTTATATATTGAATTAATATCCCTACATGTATT  
GCACATTGTAAAAAGTTTTAGTTTTGATGAGTTGTGAGTTTACCTTGTATACTGTAGGCA  
CACTTTGCACTGATATATCATGAGTGAATAAATGTCTTGCCCTACTCAAAAAAAAAAAAA

85/392

**FIGURE 82**

MDSLASLVLCGVSLLLSGTVEGAMDLILINSLPLVSDAETSLTCIASGWRPHEPITIGRD  
FEALMNQHQDPLEVTQDVTREWAKKVWVKREKASKINGAYFCEGRVRGEAIRIRTMKMRQ  
QASFLPATLTMTVDKGDVNI SFKKVLIKEEDAVIYKNGSFIHSVPRHEVPDILEVHLP  
AQPQDAGVYSARYIGGNLFTSAFTRLIVRRCEAQKWGPECNHLCTACMNNGVCHEDTGEC  
ICPPGFMGRTCEKACELHTFGRCTCKERCSCGQEGCKSYVFCLPDYPYGCSCATGWKGLQCNE  
ACHPGFYGPDCKLRSCNNGEMCDRFQGCCLSPGWQGLQCEREGIPRMTPKIVDLPDHIE  
VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAIFTIHRILPPDSG  
VWVCSVNTVAGMVEKPFENISVKVLPKPLNAPNVIDTGHNFVINISSEPYFGDGPIKSKK  
LLYKPVNHYEAWQHIQVTNEIVTLNYLEPRTEYELCVQLVRRGEGGEGHPGVRRTTAS  
IGLPPPRGLNLLPKSQTTILNLTWQPIFPSSDDFYVEVERRSVQKSDQONIKVPGNLTSV  
LLNNLHPREQYVVRARVNTKAQGEWSEDLTAWTLSDILPPQPENIKISNITHSSAVISWT  
ILDGYSISSITIRYKVQGNEDQHVDVKIKNATIIQYQLKGLEPETAYQVDIFAENNIGS  
SNPAFSHELVTLPESQAPADLGGGKMLLIAILGSAGMTCLTVLLAFLIILQLKRANVQRR  
MAQAFQNVREEPAVQFNSGTLALNRKVKNPDPTIYPVLDWNDIKFQDVIGEGNFGQVLK  
ARIKKDGLRMDAAIKRMKEYASKDDHRDFAGELEVLCKLGHPNIINLLGACEHRGYLYL  
AIEYAPHGNLLDFLRKSRVLETDPAFAIANSTASTLSSQQLLHFAADVARGMDYLSQKQF  
IHRDLAARNILVGENYVAKIADFGLSRGQEVYVKKTMGRLPVRWMAIESLNYSVYTTNSD  
VWSYGVLLWEIVSLGGTPYCGMTCAELYEKLPGYRLEKPLNCDDEVYDLMRQCWREKPY  
ERPSFAQILVSLNRMLEERKTYVNTTLYEKFTYAGIDCSAEEAA

**Signal sequence:**

1-38

**Transmembrane domain:**

750-770

**N-glycosylation site:**140-143, 158-161, 399-402, 438-441, 464-467, 560-563, 596-  
599, 649-652, 691-694, 930-933, 1011-1014, 1104-1107**cAMP- and cGMP-dependent protein kinase phosphorylation  
site:**

534-537

**Tyrosine kinase phosphorylation site:**

149-156, 808-816, 1094-1102

**N-myristoylation site:**18-23, 98-103, 187-192, 196-201, 270-275, 286-291, 295-300,  
420-425, 595-600, 984-989, 1036-1041, 1041-1046, 1115-1120**Prokaryotic membrane lipoprotein lipid attachment site:**

882-892

86/392

EGF-like domain cysteine pattern signature:

240-251, 287-298, 329-340

Tyrosine protein kinases specific active-site signature:

960-972

Protein kinase domain:

824-1092

Fibronectin type III domain:

444-529, 543-626, 639-724

EGF-like domain:

220-251, 268-298

laminin\_EGF Laminin EGF-like (Domains III and V):

219-268

Immunoglobulin domain:

156-193

Zinc finger:

295-313

Receptor tyrosine kinase:

844-868, 869-898, 936-982, 986-1024, 1025-1052, 1052-1088

87/392

**FIGURE 83**

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGCGTGCAGCAGCTCCA  
GAAAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGCCT  
GGCTCACAACAAGATGCTCAAGGTGTAGCCGTACTGTGTGTGTGTGCAGCCGCTTGGTG  
CAGTCAGTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGGGCGGTCCGACGGCGG  
TAATTTTCTGGATGATAAAACAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGG  
ACAGTGGAAACAAATTCCGAGACGAAGTAGAGGATGATTATTTCCGCACTTGGAGTCCAGG  
AAAACCCCTTCGATCAGGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAG  
TCGCCATAAAGTATGCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAG  
GCTTACACACAGGATGAAAGAAGCAGGAGTAGACCATAGGCAGTGGAGGGGTCCCATATT  
ATCCACCTGCAAGCAGTGCCAGTGGTCTATCCCAGCCCTGTTTGTGGTTTCAGATGGTCA  
TACCTACTCTTTTCAGTGCAAACTAGAAATATCAGGCATGTGTCTTAGGAAAACAGATCTC  
AGTCAAATGTGAAGGACATTGCCCATGTCTTCAGATAAGCCCACCAGTACACAGAGAAA  
TGTTAAGAGAGCATGCAGTGACCTGGAGTTCAGGGAAGTGGCAAACAGATTGCGGGACTG  
GTTCAAGGCCCTTCATGAAAGTGGAAGTCAAAACAAGAAGACAAAAACATTGCTGAGGCC  
TGAGAGAAGCAGATTCGATAACCAGCATCTTGCCAATTTGCAAGGACTCACTTGGCTGGAT  
GTTTAACAGACTTGATACAACTATGACCTGCTATTGGACCAGTCAGAGCTCAGAAGCAT  
TTACCTTGATAAGAATGAACAGTGTACCAAGGCATTCTTCAATTCTTGTGACACATACAA  
GGACAGTTTAAATATCTAATAATGAGTGGTGTACTGCTTCCAGAGACAGCAAGACCCACC  
TTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCAAGGGGTAAAGAAGCTCCTAGGACA  
GTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGCCAACACAATGTCATGGCAGTGT  
TGGACAGTGTGGTGTGTTGACAGATATGGAAATGAAGTCATGGGATCCAGAATAAATGG  
TGTTGCAGATTGTGCTATAGATTTTGAGATCTCCGGAGATTTTGCTAGTGGCGATTTTCA  
TGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATGATGAAGATGAAATTGA  
AGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTATGACCATGATGTATACAT  
TTGATTGATGACAGTTGAAATCAATAAATTCTACATTTCTAATATTTACAAAAATGATAG  
CCTATTTAAAATTATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCACATATATTTT  
GTATAATTATTTGAAAAATTGCAGCTAAAGTTATAGAACTTTATGTTTAAATAAGAATCA  
TTTGCTTTGAGTTTTTATATTCTTACACAAAAGAAAATACATATGCAGTCTAGTCAGA  
CAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTACGAGAACAACTTTGTAAAT  
CTTCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTGAAAGAT  
AATTCTAAGTGAATTTAAAATAAATAAATTTTTAATGACCTGGGTCTTAAGGATTTAGG  
AAAAATATGCATGCTTTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTC  
AGGATAACAGAGAGATACCACATGACTCCAAAAAAAAAAAAA

88/392

**FIGURE 84**

MLKVSAVLCVCAAAWCSQSLAAAAA VAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNK  
FRDEVEDDYFRTWSPGKPFQALDPAKDPCLMKCSRHKVCIAQDSQTAVCISHRRLTHR  
MKEAGVDHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCE  
GHCPCP SDKPTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKKTKTLLRPERSR  
FDTSILPICKDSLGMFNRDLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSL I  
SNNEWCYCFQRQQDPPCQTELSNIQKRQGVKLLGQYIPLCDEDGYKPTQCHGSGVGCW  
CVDRYGNEVMGSRINGVADCAIDFEISGDFASGDFHEWTDDDEDDIMNDEDEIEDDDE  
DEGDDDDGGDDHVDVYI

**Important features:****Signal peptide:**

amino acids 1-16

**Leucine zipper pattern:**

amino acids 246-267

**N-myristoylation sites:**

amino acids 357-362, 371-376 and 376-381

**Thyroglobulin type-1 repeat proteins:**

amino acids 353-365 and 339-352



89/392

**FIGURE 85**

CCCACGCGTCCGGCACTGCAGTCTCCAGCCTGAGCCATGGGCGCCGAGCCCTCCTGCTC  
CTGCTTCTGTCTTTTCTGGCGCCCTGGGCCACCATAGCCCTCCGGCCGGCCTTAAGGGCC  
CTCGGCAGCCTACACTTGCCAACCAACCCACATCCCTCCCGGCTGTAGCCAAGAACTAT  
TCGGTTCTCTACTTCCAACAGAAGGTTGATCATTTTGGATTTAATACTGTGAAAACTTTT  
AATCAGCGGTACCTAGTAGCTGATAAATACTGGAAGAAAAATGGTGGATCAATACTTTTC  
TAACTGGTAATGAAGGGGACATTATCTGGTTTTGTAAACACGGGGTTCATGTGGGAT  
GTGGCTGAGGAACTGAAAGCTATGTTGGTGTTCGCTGAACATCGATACTATGGAGAGTCT  
CTCCCCTTTGGTGACAACTCATTCAAGGATTCAGACACTTGAATTTCTGACATCAGAA  
CAAGCTCTGGCTGATTTTGCAGAGTTAATCAAACACTTGAAAAGAACAATCCCAGGAGCT  
GAAAATCAACCTGTCATTGCCATAGGAGGCTCCTATGGTGGCATGCTTGCCGCTGGTTT  
AGGATGAAATATCCTCATATGGTAGTTGGAGCTTTCGAGCTTCTGCCCCCTATCTGGCAG  
TTTGAGGATTTAGTACCTTGTGGTGTATTTATGAAGATCGTAACACTACAGATTTTAGGAA  
AGCGGTCCACATTGTTTCAGAGAGCATCCACAGGTCCTGGGATGCCATTAATCGACTCTCA  
AATACTGGCAGTGGTTTTGCAGTGGCTTACTGGAGCCCTTCACTTATGCAGCCCATTA  
TCTCAGGACATCCAACATTTGAAAGACTGGATCTCTGAAACCTGGGTGAATCTGGCAATG  
GTGGACTATCCTTATGCCTCTAACTTTTTACAGCCTTTGCCTGCTTGGCCTATCAAGGTA  
GTGTGCCAGTATTTGAAAAATCCCAATGTATCTGATTCACTGCTGCTGCAGAATATTTTC  
CAAGCTCTGAATGTATATTACAATTATTCGGGCCAGGTGAAATGCCTGAATATTTTCAGAG  
ACAGCAACTAGCAGTCTGGGAACACTGGGTTGGAGCTATCAGGCCTGCACAGAAGTAGTC  
ATGCCCTTTTGTACTAATGGTGTGATGACATGTTTGAACCTCACTCATGGAACTTAAAG  
GAACTTTCTGATGACTGTTTTCAACAGTGGGGTGTGAGACCAAGGCCCTCCTGGATCACT  
ACTATGTATGGAGGCAAAAACATTAGTTTACACACAAACATTGTTTTTTCAGCAATGGTGAA  
CTAGACCCCTGGTCAGGAGGTGGAGTAACAAAGGATATCACAGACACTCTGGTTGCAGTC  
ACCATCTCAGAGGGGGGCCACCCTTAGATCTCCGCACCAAGAATGCCTTGGATCCTATG  
TCTGTGCTGTTAGCCCGCTCCTTGGAAGTTAGACATATGAAGAATTGGATCAGAGATTTT  
TATGACAGTGCGGGAAAGCAGCACTGAGAACTTTTGATTGTTTTCAATTTCTTCTTTTA  
TGTTTACACCACCACATTCCCATTCACTTTGATTTTCTACATGTAATTACCTTCTTTTGT  
TTATCATTAGATTTGATGGGGCCAAAGTTGAGATAGAATAGAGGGTGATGACGGTAAGAG  
CAAGTGTCCCATGAATGTGATTTCTGGGTTCTCACTGTCTTTGCACCACGTCTAGGAA  
GAATCTTCTTGATAGCTCTCCACACCATCAGTGGCCCTCATAACTGGAGTAGAGTTCCCT  
GGTTGCTTTTTCATAAGAGGGAGAGTTACTTTT

90/392

**FIGURE 86**

MGRRALLLLLLSFLAPWATIALRPALRALGSLHLPTNPTSLPAVAKNYSVLYFQQKVDHF  
GFNTVKTFFNQRYLVADKYWKKNNGGSILFYTGNEGDIWFNNTGFMWDVAEELKAMLVFA  
EHYYGESLPFGDNSFKDSRHLNFLTSEQALADFAELIKHLKRTIPGAENQPVIAIGGSY  
GGMLAAWFRMKYPHVVVGALAASAPIWQFEDLVPCGVFMKIVTTDFRKSGPHCSESIHRS  
WDAINRLSNTGSGLQWLTGALHLCSPILTSQDIQHLKDWISETWVNLAMVDYPYASNFLQP  
LPAWPIKVVVCQYLKNPNVSDSLLLQNIQALNVYNYSGQVKCLNISETATSSLGTLGWS  
YQACTEVVMPFCTNGVDDMFEPHSWNLKELSDDCFQQWGVPRPSWITTMYGKNISSHT  
NIVFSNGELDPWSGGGVTKDITDTLVAVTISEGAHHLDLRTKNALDPMSVLLARSLEVRH  
MKNWIRDFYDSAGKQH

**Signal sequence:**

1-18

**Transmembrane domain:**

None

**N-glycosylation site:**

47-50, 101-104, 317-320, 336-339, 345-348, 415-418

**Glycosaminoglycan attachment site:**

433-436

**N-myristoylation site:**

178-183, 181-186, 182-187, 198-203, 339-344, 434-439

**Amidation site:**

1-4

**alpha/beta hydrolase fold:**

115-372

91/392

**FIGURE 87**

GGCGGCGTCCGTGAGGGGCTCCTTTGGGCAGGGGTAGTGTGTTGGTGTCCCTGTCTTGCGT  
GATATTGACAACTGAAGCTTTCCTGCACCACTGGACTTAAGGAAGAGTGTACTCGTAGG  
CGGACAGCTTTAGTGGCCGGCCGGCCGCTCTCATCCCCCGTAAGGAGCAGAGTCCTTTGT  
ACTGACCAAGATGAGCAACATCTACATCCAGGAGCCTCCACGAATGGGAAGGTTTATT  
GAAAACCTACAGCTGGAGATATTGACATAGAGTTGTGGTCCAAAAGACTCCTAAAGCTTG  
CAGAAATTTTATCCAACCTTGTGTTGGAAGCTTATTATGACAATACCATTTTTCATAGAGT  
TGTGCCTGGTTTCATAGTCCAAGGCGGAGATCCTACTGGCACAGGGAGTGGTGGAGAGTC  
TATCTATGGAGCGCCATTCAAAGATGAATTCATTACGGTTGCGTTTTAATCGGAGAGG  
ACTGGTTGCCATGGCAAATGCTGGTTCTCATGATAATGGCAGCCAGTTTTTCTTCACACT  
GGGTCGAGCAGATGAACCTTAACAATAAGCATACCATCTTTGGAAAGGTTACAGGGGATAC  
AGTATATAACATGTTGCGACTGTCAGAAGTAGACATTGATGATGACGAAAGACCACATAA  
TCCACACAAAATAAAAAGCTGTGAGGTTTGTGTTAATCCTTTTGATGACATCATTCCAAG  
GGAAATTAAAAGGCTGAAAAAAGAGAAACCAGAGGAGGAAGTAAAGAAATTGAAACCCAA  
AGGCACAAAAAATTTTAGTTTTACTTTTCATTTGGAGAGGAAGCTGAGGAAGAAGAGGAGGA  
AGTAAATCGAGTTAGTCAGAGCATGAAGGGCAAAAGCAAAAGTAGTCATGACTTGCTTAA  
GGATGATCCACATCTCAGTTCTGTTCCAGTTGTAGAAAGTGAAAAAGGTGATGCACCAGA  
TTTAGTTGATGATGGAGAAGATGAAAGTGACAGCATGATGAATATATTGATGGTGATGA  
AAAGAACCTGATGAGAGAAAGAATTGCCAAAAAATTAAGGACACAAAGTGCGAATGT  
TAAATCAGCTGGAGAAGGAGAAGTGGAGAAGAAATCAGTCAGCCGCAGTGAAGAGCTCAG  
AAAAGAAGCAAGACAATTAAAACGGGAACCTTTAGCAGCAAAACAAAAAAGTAGAAAA  
TGCAGCAAAACAAGCAGAAAAAAGAAGTGAAGAGGAAGAAGCCCCCTCAGATGGTGCTGT  
TGCCGAATACAGAAGAGAAAAAGCAAAAGTATGAAGCTTTGAGGAAGCAACAGTCAAAGAA  
GGGAACCTTCCCGGGAAGATCAGACCCTTGCACTGCTGAACCAGTTTAAATCTAAACTCAC  
TCAAGCAATTGCTGAAACACCTGAAAATGACATTCTGAAACAGAAGTAGAAGATGATGA  
AGGATGGATGTCACATGTACTTCAGTTTGAGGATAAAAGCAGAAAAGTGAAAGATGCAAG  
CATGCAAGACTCAGATACATTTGAAATCTATGATCCTCGGAATCCAGTGAATAAAAGAAG  
GAGGGAAGAAAGCAAAAAGCTGATGAGAGAGAAAAAAGAAAGAAGATAAAATGAGAATAA  
TGATAACCAGAACTTGCTGGAAATGTGCCTACAATGGCCTTGTAACAGCCATTGTTCCCA  
ACAGCATCACTTAGGGGTGTGAAAAGAAGTATTTTTGAACCTGTTGTCTGGTTTTGAAAA  
ACAATTATCTTGTGTTTGCAAATTGTGGAATGATGTAAGCAAATGCTTTTGGTTACTGGTA  
CATGTGTTTTTCTAGCTGACCTTTTATATTGCTAAATCTGAAATAAAATAACTTTTCT  
TCCACAAAAA

92/392

**FIGURE 88**

MSNIYIQEPPTNGKVLLKTTAGDIDIELWSKEAPKACRNFIQLCLEAYYDNTIFHRVVP  
FIVQGGDPTGTGSGGESIYGAPFKDEFHSRLRFNRRGLVAMANAGSHDNGSQFFFTLGRA  
DELNNKHTIFGKVTGDTVYNMLRLSEVDIDDDERPHNPHKIKSCEVLFPFDDIIPREIK  
RLKKEKPEEEVKKLKPKGTKNFSLLSFGBEEAEEEEEVNRVSQSMKGKSKSSHDLKDDP  
HLSSVPVVESEKGDAPDLVDDGEDESAEHDEYIDGDEKNLMRERIAKKLKKDTSANVKS  
A  
GEGEVEKKSVSRSEELRKEARQLKRELLAAKQKKVENAAKQAEKRSEEEEAPPDGAVAEY  
RREKQKYEALRKQQSKKGTSRDQTLALLNQFKSKLTQAIAETPENDIPETEVEDDEGWM  
SHVLQFEDKSRKVKDASMQSDTFEIIDPRNPFVNKRRREESKKLMREKKERR

**Important features:****Signal peptide:**

amino acids 1-21

**N-glycosylation sites:**

amino acids 109-112 and 201-204

**Cyclophilin-type peptidyl-prolyl cis-trans isomerase  
signature:**

amino acids 49-66

**Homologous region to Cyclophilin-type peptidyl-prolyl cis-  
trans isomerase:**

amino acids 96-140, 49-89 and 22-51

93/392

**FIGURE 89**

CCCGGCTCCGCTCCCTCTGCCCCCTCGGGGTCGCGCGCCACGATGCTGCAGGGCCCTGG  
CTCGCTGCTGCTGCTCTTCCTCGCCTCGCACTGCTGCCTGGGCTCGGCGCGGGGCTCTT  
CCTCTTTGGCCAGCCCGACTTCTCCTACAAGCGCAGCAATTGCAAGCCCATCCCGGTCAA  
CCTGCAGCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCAACCTGCTGGGCCA  
CGAGACCATGAAGGAGGTGCTGGAGCAGGCCGGCGCTTGGAATCCCGCTGGTCATGAAGCA  
GTGCCACCCGGACACCAAGAAGTTCTGTGCTCGCTCTTCGCCCCCGTCTGCCTCGATGA  
CCTAGACGAGACCATCCAGCCATGCCACTCGCTCTGCGTGCAAGGACCGCTGCGC  
CCCGGTCTGTCCGCCTTCGGCTTCCCCCTGGCCCGACATGCTTGAGTGCGACCGTTTCCC  
CCAGGACAACGACCTTTGCATCCCCCTCGCTAGCAGCGACCACCTCCTGCCAGCCACCGA  
GGAAGCTCCAAAGGTATGTGAAGCCTGCAAAAATAAAAATGATGATGACAACGACATAAT  
GGAAACGCTTTGTAAAAATGATTTTGCCTGAAAATAAAAGTGAAGGAGATAACCTACAT  
CAACCGAGATACCAAAATCATCCTGGAGACCAAGAGCAAGACCATTTACAAGCTGAACGG  
TGTGTCCGAAAGGGACCTGAAGAAATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGCAC  
CTGTGAGGAGATGAACGACATCAACGCGCCCTATCTGGTCATGGGACAGAAACAGGGTGG  
GGAGCTGGTGATCACCTCGGTGAAGCGGTGGCAGAAGGGGCAGAGAGAGTTCAAGCGCAT  
CTCCCGCAGCATCCGCAAGCTGCAGTGCTAGTCCCGGCATCCTGATGGCTCCGACAGGCC  
TGCTCCAGAGCACGGCTGACCATTTCTGCTCCGGGATCTCAGCTCCCGTTCCCCAAGCAC  
ACTCCTAGCTGCTCCAGTCTCAGCCTGGGCAGCTTCCCCCTGCCTTTTGCACGTTTGCAT  
CCCCAGCATTTCTGAGTTATAAGGCCACAGGAGTGGATAGCTGTTTTACCTAAAGGAA  
AAGCCCACCCGAATCTTGTAGAAATATTCAAACATAATAAATCATGAATATTTTAA

94/392

## FIGURE 90

MLQGPGLLLLLFLASHCCCLGSARGLFLEFGQPDFS YKRSNCKPIPVNLQLCHGIEYQNMRL  
PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDDETIQPCHS LCVQ  
VKDRCAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKND  
DDNDIMETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLNGVSE RDLKKSVLWLK  
DSLQCTCEEMNDINAPYLVMGQKQGGELVITSVKRWQKGQREFKRISR SIRKLQC

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homologous to frizzled N terminus:

amino acids 6-153

GGAAGGGAGGAGCAGGCCACACAGGCACAGGCCGGTGAGGGACCTGCCCAGACCTGGAG  
 GGTCTCGCTCTGTGCACAGGCTGGAGTGCAGTGGTGTGATCTTGGCTCATCGTAACCTC  
 CACCTCCCGGGTTCAAGTGATTCTCATGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGT  
 GGTGACTTCCAAGAGTGACTCCGTCGGAGGAAAATGACTCCCCAGTCGCTGCTGCAGACG  
 ACACTGTTTCTGCTGAGTCTGCTCTTCTGGTCCAAGGTGCCACGGCAGGGGCCACAGG  
 GAAGACTTTTCGCTTCTGCAGCCAGCGGAACCAGACACACAGGAGCAGCCTCCACTACAAA  
 CCCACACCAGACCTGCGCATCTCCATCGAAGACTCCGAAGAGGCCCTCACAGTCCATGCC  
 CCTTTCCTGTCAGCCCCACCCTGCTTCCCGACTCCTTCCCTGACCCAGGGGCCCTTACCA  
 TTCTGCCTCTACTGGAACCGACATGCTGGGAGATTACATCTTCTCTATGGCAAGCGTGAC  
 TTCTTGCTGAGTGACAAAGCCTCTAGCCTCCTCTGCTTCCAGCACCAGGAGGAGAGCCTG  
 GCTCAGGGCCCCCGCTGTTAGCCACTTCTGTACCTCTCTGGTGAGGCCCTCAGAACATC  
 AGCCTGCCCAGTGCCGCCAGCTTACCTTCTCCTTCCACAGTCTCTCCACACGGCCCGCT  
 CACAATGCCTCGGTGGACATGTGCGAGCTCAAAGGGACCTCCAGCTGCTCAGCCAGTTT  
 CTGAAGCATCCCCAGAAGGCCTCAAGGAGGCCCTCGGCTGCCCCCGCCAGCCAGCAGTTG  
 CAGAGCCTGGAGTCGAAACTGACCTCTGTGAGATTATGGGGGACATGGTGTCTTTCGAG  
 GAGGACCGGATCAACGCCACGGTGTGGAAGCTCCAGCCCACAGCCGGCCTCCAGGACCTG  
 CACATCCACTCCCGGCAGGAGGAGGAGCAGAGCGAGATCATGGAGTACTCGGTGCTGCTG  
 CCTCGAACACTCTTCCAGAGGACGAAAGGCCGGAGCGGGGAGGCTGAGAAGAGACTCCTC  
 CTGGTGGACTTTCAGCAGCCAAGCCCTGTTCCAGGACAAGAATTCCAGCCAAGTCTGGGT  
 GAGAAGGTCTTGGGGATTGTGGTACAGAACACCAAAGTAGCCAACCTCACGGAGCCCGTG  
 GTGCTCACTTTCAGCACCAGCTACAGCCGAAGAATGTGACTCTGCAATGTGTGTTCTGG  
 GTTGAAGACCCCAATTGAGCAGCCCCGGGGCATTGGAGCAGTGCTGGGTGTGAGACCGTC  
 AGGAGAGAAACCCAAACATCCTGCTTCTGCAACCACTTGACCTACTTTGCAGTGCTGATG  
 GTCTCCTCGGTGGAGGTGGACGCCGTGCACAAGCACTACCTGAGCCTCCTCTCCTACGTG  
 GGCTGTGTCTCTCTGCCCTGGCCTGCCTTGTCAACATTGCCGCCTACCTCTGCTCCAGG  
 GTGCCCCCTGCCGTGCAGGAGGAAACCTCGGGACTACACCATCAAGGTGCACATGAACCTG  
 CTGCTGGCCGTCTTCTGCTGGACACGAGCTTCTGCTCAGCGAGCCGGTGGCCCTGACA  
 GGCTCTGAGGCTGGCTGCCGAGCCAGTGCCATCTTCTGCACTTCTCCCTGCTCACCTGC  
 CTTTCTGGATGGCCTCGAGGGGTACAACCTCTACCGACTCGTGGTGGAGGTCTTTGGC  
 ACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGGCTTCCCCATCTTT  
 CTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAACCTATGGCCCCATCATCTTGGCTGTG  
 CATAGGACTCCAGAGGGCGTCATCTACCTTCCATGTGCTGGATCCGGGACTCCCTGGTC  
 AGCTACATACCAACCTGGGCCTCTTACGCTGGTGTCTTCTGTTCAACATGGCCATGCTA  
 GCCACCATGGTGGTGCAGATCCTGCGGCTGCGCCCCACACCCAAAAGTGGTCAATGTG  
 CTGACACTGCTGGGCCTCAGCCTGGTCTTGGCCTGCCCTGGGCCCTGATCTTCTTCTCC  
 TTTGCTTCTGGCACCTTCCAGCTTGTGCTCCTCTACCTTTTTCAGCATCATCACCTCCTTC  
 CAAGGCTTCTCATCTTCTATCTGGTACTGGTCCATGCGGCTGCAGGCCCGGGGTGGCCCC  
 TCCCCCTCTGAAGAGCAACTCAGACAGCGCCAGGCTCCCCATCAGCTCGGGCAGCACCTCG  
 TCCAGCCGCATCTAGGCCTCCAGCCCACCTGCCCATGTGATGAAGCAGAGATGCGGCCCTC  
 GTCGCACACTGCCTGTGGCCCCCGAGCCAGGCCAGCCCCAGGCCAGTCAGCCGCAGACT  
 TTGGAAGCCCAACGACCATGGAGAGATGGGCGCTTGGCATGGTGGACGAGCTCCCGGGC  
 TGGGCTTTTGAATTGGCCTTGGGGACTACTCGGCTCTCACTCAGCTCCACGGGACTCAG  
 AAGTGCGCCGCCATGCTGCCTAGGGTACTGTCCCCACATCTGTCCCAACCCAGCTGGAGG  
 CCTGGTCTCTCCTTACAACCCCTGGGCCAGCCCTCATTGCTGGGGGCCAGGCCCTTGGAT  
 CTTGAGGGTCTGGCACATCCTTAATCCTGTGCCCTGCCTGGGACAGAAATGTGGCTCCA  
 GTTGCTCTGTCTCTCGTGGTCAACCTGAGGGCACTCTGCATCCTCTGTCAATTTTAACCTC  
 AGGTGGCACCCAGGGCGAATGGGGCCCAGGGCAGACCTTCAGGGCCAGAGCCCTGGCGGA





97/392

**FIGURE 92**

MTPQSLQLQTTLFLLSLLFLVQGAHGRGHREDFRFCSQRNQTHRSSLHYKPTPDLRISIEN  
SEEALTVHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFFLLSDKASSLL  
CFQHQEESLAQGPPLLATSVTSWWSPQNISLPSAASFSTFSFHSPHTAAHNASVDMCELK  
RDLQLLSQFLKHPQKASRRPSAAPASQQLSLESKLTSVRFMGDMVSFEDRINATVWKL  
QPTAGLQDLHIHSRQEEEQSEIMEYSVLLPRTLTFQRTKGRSGEAEKRLLLVDFSSQALFQ  
DKNSSQVLGEKVLGIVVQNTKVANLTEPVVLTQHQQLQPKNVTLCVFWVEDPTLSSPGH  
WSSAGCETVRRETQTSCFCNHLTYFAVLMVSSVEVDAVHKHYLSLLSYVGCVVVSALACLIV  
TIAAYLCSRVPPLPCRKPRDYTIKVHMNLLLAFLLDTSFLLSEPVALTGSEAGCRASAI  
FLHFSLLTCLSWMGLEGYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVTLVALVDVD  
NYGPIILAVHRTPEGVIYPSMCWIRDSLVSYITNLGLFSLVFLFNMAMLATMVVQILRLR  
PHTQKWSHVLTLGLSLVLGLPWALIFFSFASGTFQLVVLYLFSIITSFQGFLIFIWYWS  
MRLQARGGPSPLKSNSDSARLPISSGSTSSRI

**Important features:****Signal peptide:**

amino acids 1-25

**Putative transmembrane domains:**amino acids 382-398, 402-420, 445-468, 473-491, 519-537,  
568-590 and 634-657**Microbodies C-terminal targeting signal:**

amino acids 691-693

**cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

amino acids 198-201 and 370-373

**N-glycosylation sites:**amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-  
327 and 341-344**G-protein coupled receptors family 2 proteins:**

amino acids 475-504

98/392

**FIGURE 93**

CCCACGCGTCCGAAGGCAGACAAAGGTTCAATTTGTAAAGAAGCTCCTTCCAGCACCTCCT  
CTCTTCTCCTTTTGCCCAAACACCCAGTGAGTGTGAGCATTTAAGAAGCATCCTCTGC  
CAAGACCAAAGGAAAGAAGAAAAAGGGCCAAAGCCAAATGAAACTGATGGTACTTGT  
TTTCACCATTTGGGCTAACTTTGCTGCTAGGAGTTCAAGCCATGCCTGCAAATCGCCTCTC  
TTGCTACAGAAAGATACTAAAAGATCACAACTGTCACAACCTTCCGGAAGGAGTAGCTGA  
CCTGACACAGATTGATGTCAATGTCCAGGATCATTTCTGGGATGGGAAGGGATGTGAGAT  
GATCTGTTACTGCAACTTCAGCGAATTGCTCTGCTGCCCAAAGACGTTTTCTTTGGACC  
AAAGATCTCTTTTCGTGATTCCTTGCAACAATCAATGAGAATCTTCATGTATTCTGGAGAA  
CACCATTCCTGATTTCCACAAACTGCACTACATCAGTATAACTGCATTTCTAGTTTCTA  
TATAGTGCAATAGAGCATAGATTCTATAAATTCTTACTTGTCTAAGACAAGTAAATCTGT  
GTTAAACAAGTAGTAATAAAAGTTAATTCATCTAAAAAAAAAAAAA

99/392

## FIGURE 94

MKLMVLVFTIGLTLLLGVQAMPANRLSCYRKILKDHNCHNLPEGVADLTQIDVNVQDHFWDGKGCEMICYCNFSELLCCPKDVFFGPKISFVIPCNNQ

**Important features:**

**Signal peptide:**

amino acids 1-20

**N-glycosylation site:**

amino acids 72-76

**Tyrosine kinase phosphorylation site:**

amino acids 63-71

100/392

**FIGURE 95**

GAATTCGGGGCCCCAGGATGCCAACTTTGAATAGGATGAAGACTACAACCTTGTTCCCTTC  
TCATCTGCATCTCCCTGCTCCAGCTGATGGTCCCAGTGAATACTGATGAGACCATAGAGA  
TTATCGTGGAGAATAAGGTCAAGGAACTTCTTGCCAATCCAGCTAACTATCCCTCCACTG  
TAACGAAGACTCTCTCTTGCACTAGTGTCAAGACTATGAACAGATGGGCCTCCTGCCCTG  
CTGGGATGACTGCTACTGGGTGTGCTTGTGGCTTTGCCTGTGGATCTTGGGAGATCCAGA  
GTGGAGATACTTGCAACTGCCTGTGCTTACTCGTTGACTGGACCACTGCCCCTGCTGCC  
AACTGTCCTAAGAATGAAGAGGTGGAGAACCCAGCTTTGATATGATGAATCTAACAAAA  
CTGCAGTCTCAATTTGGAAATCTGACTCATGTGCCTTTAAATGTGTTTCATATTGCCCAT  
TACCCTGCTTCTTGAAATGCTTCTTGAAAAATAAAGACAAATTTGCATGTG

101/392

## FIGURE 96

MKTTTCSLLICISLLQLMVPVNTDETIEIIVENKVKELLANPANYPSTVTKTSLCTSVK  
TMNRWASCPAGMTATGCACGFACGSWEIQSGDTCNCLCLLDWTTARCCQLS

102/392

**FIGURE 97**

GAGGCAGAAAGGCAGAAAGGAGAAAATTTCAGGATAACTCTCCTGAGGGGTGAGCCAAGCC  
CTGCCATGTAGTGCACGCAGGACATCAACAAACACAGATAACAGGAAATGATCCATTCCC  
TGTGGTCACTTATTCTAAAGGCCCAACCTTCAAAGTTCAAGTAGTGATATGGATGACTC  
CACAGAAAGGGAGCAGTCACGCCTTACTTCTTGCCCTTAAGAAAAGAGAAGAAATGAAACT  
GAAGGAGTGTGTTTCCATCCTCCCACGGAAGGAAAGCCCCTCTGTCCGATCCTCCAAAGA  
CGGAAAGCTGCTGGCTGCAACCTTGCTGCTGGCACTGCTGTCTTGCTGCCTCACGGTGGT  
GTCTTTCTACCAGGTGGCCGCCCTGCAAGGGGACCTGGCCAGCCTCCGGGCAGAGCTGCA  
GGGCCACCACGCGGAGAAGCTGCCAGCAGGAGCAGGAGCCCCAAGGCCGGCCTGGAGGA  
AGCTCCAGCTGTCACCGCGGACTGAAAATCTTTGAACCACCAGCTCCAGGAGAAGGCAA  
CTCCAGTCAGAACAGCAGAAATAAGCGTGCCGTTTCAGGGTCCAGAAGAAACAGTCACTCA  
AGACTGCTTGCAACTGATTGCAGACAGTGAACACCAACTATACAAAAAGGATCTTACAC  
ATTTGTTCCATGGCTTCTCAGCTTTAAAAGGGGAAGTGCCCTAGAAGAAAAAGAGAATAA  
AATATTGGTCAAAGAAACTGGTTACTTTTTTATATATGGTCAGGTTTATATACTGATAA  
GACCTACGCCATGGGACATCTAATTCAGAGGAAGAAGGTCCATGTCTTTGGGGATGAATT  
GAGTCTGGTGACTTTGTTTCGATGTATTCAAAATATGCCTGAAACACTACCCAATAATTC  
CTGCTATTTCAGCTGGCATTGCAAACTGGAAGAAGGAGATGAACTCCAATTGCAATACC  
AAGAGAAAATGCACAAATATCACTGGATGGAGATGTCACATTTTTTGGTGCATTGAAACT  
GCTGTGAACCTACTTACACCATGTCTGTAGCTATTTTCCTCCCTTTCTCTGTACCTCTAAG  
AAGAAAGAATCTAACTGAAAATACCAAAAAAAAAAAAAAAAAA

103/392

**FIGURE 98**

MDDSTEREQSRLTSCCLKKREEMKLKECVSILPRKESPSVRSSKDGKLLAATLLLALLSCC  
LTVVSFYQVAALQGD LASLRAELQGHHA EKL PAGAGAPKAGLEEAPAVTAGLKI FEPPAP  
GEGNSSQNSRNKRAVQGPEETVTQDCLQLIADSETPTIQKGSYTFVPWLLSFKRGSAL EE  
KENKILVKETGYFFIYGQVLYTDKTYAMGH LIQRKKVHVFGDELSLVTLFR CIQNMPETL  
PNNSCYSAGIAKLEEGDELQLAIPRENAQISLDGDVTFFGALKLL

**Transmembrane domain:**

amino acids 47-72

**N-glycosylation site:**

amino acids 124-127, 242-245

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

amino acids 33-36, 173-176

**N-myristoylation site:**

amino acids 96-101

**TNF family proteins:**

amino acids 172-206

104/392

**FIGURE 99**

CCGAGGTTGGCGATCGCTGAGAGGCAGGAGGGCCGAGGCGGGCCTGGGAGGCGGCCGGAG  
GTGGGGCGCCGCTGGGGCCGGCCCGCACGGGCTTCATCTGAGGGCGCACGGCCCGCGACC  
GAGCGTGCGGACTGGCCTCCCAAGCGTGGGGCGACAAGCTGCCGGAGCTGCAATGGGGCCG  
CGGCTGGGGATTCTTGTTTGGCCTCCTGGGCGCCGTGTGGCTGCTCAGCTCGGGCCACGG  
AGAGGAGCAGCCCCCGGAGACAGCGGCACAGAGGTGCTTCTGCCAGGTTAGTGGTTACTT  
GGATGATTGTACCTGTGATGTTGAAACCATTGATAGATTTAATAACTACAGGCTTTTCC  
AAGACTACAAAACTTCTTGAAAGTGACTACTTTAGGTATTACAAGGTAAACCTGAAGAG  
GCCGTGTCCTTTCTGGAATGACATCAGCCAGTGTGGAAGAAGGGACTGTGCTGTCAAACC  
ATGTCAATCTGATGAAGTTCCTGATGGAATTAAATCTGCGAGCTACAAGTATTCTGAAGA  
AGCCAATAATCTCATTGAAGAATGTGAACAAGCTGAACGACTTGGAGCAGTGGATGAATC  
TCTGAGTGAGGAAACACAGAAGGCTGTTCTTCAGTGGACCAAGCATGATGATTCTTCAGA  
TAACTTCTGTGAAGCTGATGACATTGAGTCCCTGAAGCTGAATATGTAGATTTGCTTCT  
TAATCCTGAGCGCTACACTGGTTACAAGGGACCAGATGCTTGGAAAATATGGAATGTGAT  
CTACGAAGAAAACCTGTTTTAAGCCACAGACAATTAAAGACCTTTAAATCCTTTGGCTTC  
TGGTCAAGGGACAAGTGAAGAGAACACTTTTTACAGTTGGCTAGAAGGTCTCTGTGTAGA  
AAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAAATGTGCATTTGAG  
TGCAAGATATCTTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACAACATTACAGA  
ATTTCAACAGCGATTTGATGGAATTTTGACTGAAGGAGAAGGTCCAAGAAGGCTTAAGAA  
CTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTTATCCAAAGTGTTACCATTCTTCGA  
GCGCCCGAGATTTTCAACTCTTTACTGGAATAAAATTCAGGATGAGGAAAACAAAATGTT  
ACTTCTGGAATACTTTCATGAAATCAAGTCATTTCCCTTTGCATTTTGATGAGAATTCATT  
TTTTGCTGGGGATAAAAAAGAAGCACACAACTAAAGGAGGACTTTGACTGCATTTTAG  
AAATATTTCAAGAATTATGGATTGTGTTGGTTGTTTTAAATGTGCTGTGTGGGGAAAGCT  
TCAGACTCAGGGTTTGGGCACTGCTCTGAAGATCTTATTTTCTGAGAAATTGATAGCAAA  
TATGCCAGAAAGTGGACCTAGTTATGAATTCATCTAACCAGACAAGAAATAGTATCATT  
ATTCAACGCATTTGGAAGAATTTCTACAAGTGTGAAAGAATTAGAAAACTTCAGGAACCTT  
GTTACAGAAATATTCATTAAAGAAAAACAAGCTGATATGTGCCTGTTTCTGGACAATGGAGG  
CGAAAGAGTGGAAATTTCTATTCAAAGGCATAATAGCAATGACAGTCTTAAGCCAAACATTT  
TATATAAAGTTGCTTTTGTAAAGGAGAATTATATTGTTTTAAGTAAACACATTTTTAAAA  
ATTGTGTTAAGTCTATGTATAATACTACTGTGAGTAAAGTAATACTTTAATAATGTGGT  
ACAAATTTTTAAGTTTTAATATTGAATAAAAGGAGGATTATCAAATTAATAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAA



105/392

**FIGURE 100**

MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAARCFQCQVSGYLDDCTCDVETIDRFNNYR  
LFPRLQKLLESDFRYYKVNLRPCPFWNDISQCGRRDCAVKPCQSDEVDPDGIKSASYKY  
SEEANNLIEECEQAERLGAVDESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVD  
LLNPERYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGL  
CVEKRAFYRLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRR  
LKNLYFLYLIELRALSKVLPFFERPDFQLFTGNKIQDEENKMLLLEILHEIKSFPLHFDE  
NSFFAGDKKEAHKLKEDFRLHFRNISRIMDCVGCFCRLWGKLQTQGLGTALKILFSEKL  
IANMPESGPSYEFHLTRQEIVSLFNAFGRISTSVKELENFRNLLQNIH

**Important features:**

**Signal peptide:**

amino acids 1-23

**N-glycosylation site:**

amino acids 280-283 and 384-387

**Amidation site:**

amino acids 94-97

**Glycosaminoglycan attachment site:**

amino acids 20-23 and 223-226

**Aminotransferases class-V pyridoxal-phosphate:**

amino acids 216-222

**Interleukin-7 proteins:**

amino acids 338-343

106/392

**FIGURE 101**

GCCTAGCCAGGCCAAGAATGCAATTGCCCCGGTGGTGGGAGCTGGGAGACCCCTGTGCTT  
GGACGGGACAGGGTCGGGGGACACGCAGGATGAGCCCCGCGACCACTGGGCACATTCTTGC  
TGACAGTGACAGTATTTTCTCCAAGGTACACTCCGATCGGAATGTATACCCATCAGCAG  
GTGTCCTCTTTGTTTCATGTTTTGGAAAGAGAATATTTTAAGGGGGAATTTCCACCTTACC  
CAAAACCTGGCGAGATTAGTAATGATCCCATAACATTTAATACAAATTTAATGGGTACC  
CAGACCGACCTGGATGGCTTCGATATATCCAAAGGACACCATATAGTGATGGAGTCCTAT  
ATGGGTCCCCAACAGCTGAAAATGTGGGGAAGCCAACAATCATTGAGATAACTGCCTACA  
ACAGGCGCACCTTTGAGACTGCAAGGCATAATTTGATAATTAATATAATGTCTGCAGAAG  
ACTTCCCGTTGCCATATCAAGCAGAATTCTTCATTAAGAATATGAATGTAGAAGAAATGT  
TGGCCAGTGAGGTTCTTGGAGACTTTCTTGGCGCAGTGAAAAATGTGTGGCAGCCAGAGC  
GCCTGAACGCCATAAACATCACATCGGCCCTAGACAGGGGTGGCAGGGTGGCACTTCCCA  
TTAATGACCTGAAGGAGGGCGTTTATGTCATGGTTGGTGCAGATGTCCCGTTTTCTTCTT  
GTTTACGAGAAGTTGAAAATCCACAGAATCAATTGAGATGTAGTCAAGAAATGGAGCCTG  
TAATAACATGTGATAAAAAATTTCTGACTCAATTTTACATTGACTGGTGCAAAATTTTCAT  
TGTTTGATAAAACAAAGCAAGTGTCCACCTATCAGGAAGTGATTCGTGGAGAGGGGATTT  
TACCTGATGGTGGAGAATACAAACCCCTTCTGATTCTTTGAAAAGCAGAGACTATTACA  
CGGATTTTCTAATTACACTGGCTGTGCCCTCGGCAGTGGCACTGGTCCTTTTTCTAATAC  
TTGCTTATATCATGTGCTGCCGACGGGAAGGCGTGAAAAGAGAAACATGCAACACCAG  
ACATCCAACCTGGTCCATCACAGTGCTATTTCAGAAATCTACCAAGGAGCTTCGAGACATGT  
CCAAGAATAGAGAGATAGCATGGCCCCCTGTCAACGCTTCCTGTGTTCCACCCTGTGACTG  
GGGAAATCATACCTCCTTTACACACAGACAACTATGATAGCACAAACATGCCATTGATGC  
AAACGCAGCAGAACTTGCCACATCAGACTCAGATTCCCCAACAGCAGACTACAGGTAAAT  
GGTATCCCTGAAGAAAGAAAATGACTGAAGCAATGAATTTATAATCAGACAATATAGCA  
GTTACATCACATTTCTTTTCTTCCAATAATGCATGAGCTTTTCTGGCATATGTTATGC  
ATGTTGGCAGTATTAAGTGTATACCAATAATACAACTTTTCAATTTTACTAATGTA  
TTTTTTTGTAATAAGCATTTTGTACAATTTGTAAACATTGATGACTTTATATTGTT  
ACAATAAAAGTTGATCTTTAAAATAAATATTATTAATGAAGCCTAAAAAAAAAAAA

107/392

**FIGURE 102**

MQLPRWWE LGDPCAWTGQGRGTRRMSPATTGTFLLT VYSIFSKVHSDRNVYPSAGVLFVH  
VLEREYFKGEFPPYPKPGEISNDPITFNTNLMGYPD RPGWLRYIQRTPYSDGVLYGSPTA  
ENVGKPTII EITAYNRRTFETARHNLIINIMSAEDFPLPYQAEFFIKNMNVEEMLASEVL  
GDFLGAVKNVWQPERLNAINITSALDRGGRVPLPINDLKEGVYVMVGADV PFSSCLREVE  
NPQNQLRCSQEMEPVITCDKKFRTQFYIDWCKISLV DTKQVSTYQEVIRGEGILPDGGE  
YKPPSDLKSRDYYTDFLITLAVPSAVALVFLILAYIMCCRREGVEKRNMQTPDIQLVH  
HSAIQKSTKELRDMSKNREIAWPLSTLPVFHPVTGEIIPPLHTDNYDSTNMPLMQTQQNL  
PHQTQIPQQOTTGKWYP

signal sequence:

Amino acids 1-46

transmembrane domain:

Amino acids 319-338

N-glycosylation site:

Amino acids 200-204

cAMP- and cGMP-dependent protein kinase phosphorylation  
site:

Amino acids 23-27

Tyrosine kinase phosphorylation site:

Amino acids 43-52

N-myristoylation sites:

Amino acids 17-23;112-118;116-122;185-191

108/392

**FIGURE 103**

CAGAAGAGGGGGCTAGCTAGCTGTCTCTGCGGACCAGGGAGACCCCCGCGCCCCCCCCGGT  
GTGAGGCGGCCTCACAGGGCCGGGTGGGCTGGCGAGCCGACGCGGCGGCGGAGGAGGCTG  
TGAGGAGTGTGTGGAACAGGACCCGGGACAGAGGAACCATGGCTCCGCAGAACCTGAGCA  
CCTTTTGCCTGTTGCTGCTATACCTCATCGGGGCGGTGATTGCCGGACGAGATTTCTATA  
AGATCTTGGGGGTGCCTCGAAGTGCCTCTATAAAGGATATTAAAAAGGCCTATAGGAAAC  
TAGCCCTGCAGCTTCATCCCGACCGGAACCCTGATGATCCACAAGCCCAGGAGAAATTCC  
AGGATCTGGGTGCTGCTTATGAGGTTCTGTCTAGATAGTGAGAAACGGAAACAGTACGATA  
CTTATGGTGAAGAAGGATTAAAAGATGGTCATCAGAGCTCCCATGGAGACATTTTTTTCAC  
ACTTCTTTGGGGATTTTGGTTTCATGTTTGGAGGAACCCCTCGTCAGCAAGACAGAAATA  
TTCCAAGAGGAAGTGATATTATTGTAGATCTAGAAGTCACTTTGGAAGAAGTATATGAG  
GAAATTTTGTGGAAGTAGTTAGAAACAAACCTGTGGCAAGGCAGGCTCCTGGCAACCGGA  
AGTGCAATTGTTCGGCAAGAGATGCGGACCACCCAGCTGGGCCCCTGGGCGCTTCCAAATGA  
CCCAGGAGGTGGTCTGCGACGAATGCCCTAATGTCAAACCTAGTGAATGAAGAACGAACGC  
TGGAAGTAGAAATAGAGCCTGGGGTGAGAGACGGCATGGAGTACCCCTTTATTGGAGAAG  
GTGAGCCTCACGTGGATGGGGAGCCTGGAGATTTACGGTTCCGAATCAAAGTTGTCAAGC  
ACCCAATATTTGAAAGGAGAGGAGATGATTTGTACACAAATGTGACAATCTCATTAGTTG  
AGTCACTGGTTGGCTTTGAGATGGATATTACTCACTTGGATGGTCACAAGGTACATATTT  
CCCGGATAAGATCACCAGGCCAGGAGCGAAGCTATGGAAGAAAGGGGAAGGGCTCCCCA  
ACTTTGACAACAACAATATCAAGGGCTCTTTGATAATCACTTTTGATGTGGATTTTCCAA  
AAGAACAGTTAACAGAGGAAGCGAGAGAAGGTATCAAACAGCTACTGAAACAAGGGTCAG  
TGCAGAAGGTATACAATGGACTGCAAGGATATTTGAGAGTGAATAAAATTTGGACTTTGTTT  
AAAATAAGTGAATAAGCGATATTTATTATCTGCAAGGTTTTTTTTGTGTGTGTTTTTGT  
TTATTTTCAATATGCAAGTTAGGCTTAATTTTTTTTATCTAATGATCATCATGAAATGAAT  
AAGAGGGCTTAAGAATTTGTCCATTTGCATTTCGGAAAAGAATGACCAGCAAAAGGTTTAC  
TAATACCTCTCCCTTTGGGGATTTAATGTCTGGTGCTGCCGCCTGAGTTTCAAGAATTAA  
AGCTGCAAGAGGACTCCAGGAGCAAAAAGAAACACAATATAGAGGGTTGGAGTTGTTAGCA  
ATTTTCATTCAAATGCCAACTGGAGAAGTCTGTTTTTAAATACATTTTGTGTTATTTTTTA

109/392

**FIGURE 104**

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDD  
PQAQEKFQDLGAAYEVLSDSEKQYDQTYGEEGLKDGHQSSHGDI FSHFFGDFGFMFGGT  
PRQQRNIPRGSDIIVDLEVTLEEYAGNFVEVVRNKPVARQAPGKRKCNCRQEMRTTQL  
GPGRFQMTQEVVCDPCPNVCLVNEERTLEVEIEPGVRDGMETPFGEPEPHVDGEPGDLR  
FRIKVVKHPIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVVHISRDKITRPGAKLW  
KKGEGLPNFDNNNIKGSLIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY

**Important features:**

**Signal peptide:**

amino acids 1-22

**Cell attachment sequence:**

amino acids 254-257

**Nt-dnaJ domain signature:**

amino acids 67-87

**Homologous region to Nt-dnaJ domain proteins:**

amino acids 26-58

**N-glycosylation site:**

amino acids 5-9, 261-265

**Tyrosine kinase phosphorylation site:**

amino acids 253-260

**N-myristoylation site:**

amino acids 18-24, 31-37, 93-99, 215-221

**Amidation site:**

amino acids 164-168

110/392

**FIGURE 105**

GGCACGAGGCGGCGGGGCAGTCGCGGGATGCGCCCCGGGAGCCACAGCCTGAGGCCCTCAG  
GTCTCTGCAGGTGTCTGTGGAGGAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACT  
TCCAGCAGCTTTAGCCCATGAGGAGGATGTGACCGGGACTGAGTCAGGAGCCCTCTGGAA  
GCATGGAGACTGTGGTGATTGTTGCCATAGGTGTGCTGGCCACCATCTTTCTGGCTTCGT  
TTGCAGCCTTGCTGCTGGTTTGCAGGCAGCGCTACTGCCGCGCGGAGACCTGCTGCAGC  
GCTATGATTCTAAGCCCATTTGTGGACCTCATTTGGTGCCATGGAGACCCAGTCTGAGCCCT  
CTGAGTTAGAACTGGACGATGTCGTTATCACCAACCCCCACATTGAGGCCATTCTGGAGA  
ATGAAGACTGGATCGAAGATGCCTCGGGTCTCATGTCCCACTGCATTGCCATCTTGAAGA  
TTTGTCACTCTGACAGAGAAGCTTGTTGCCATGACAATGGGCTCTGGGGCCAAGATGA  
AGACTTCAGCCAGTGTGACGACATCATTTGTGGTGGCCAAGCGGATCAGCCCCAGGGTGG  
ATGATGTTGTGAAGTCGATGTACCCTCCGTTGGACCCCAAACCTCCTGGACGCACGGACGA  
CTGCCCTGCTCCTGTCTGTGTCAGTCACCTGGTGCTGGTGACAAGGAATGCCTGCCATCTGA  
CGGGAGGCCTGGACTGGATTGACCAGTCTCTGTGCTGGCTGCTGAGGAGCATTTGGAAGTCC  
TTCGAGAAGCAGCCCTAGCTTCTGAGCCAGATAAAGGCCTCCAGGCCCTGAAGGCTTCC  
TGCAGGAGCAGTCTGCAATTTAGTGCCTACAGGCCAGCAGCTAGCCATGAAGGCCCTGC  
CGCCATCCCTGGATGGCTCAGCTTAGCCTTCTACTTTTTCTATAGAGTTAGTTGTTCTC  
CACGGCTGGAGAGTTGAGCTGTGTGTGCATAGTAAAGCAGGAGATCCCCGTCAGTTTATG  
CCTCTTTTGCAGTTGCAAACTGTGGCTGGTGAGTGGCAGTCTAATACTACAGTTAGGGGA  
GATGCCATTCACTCTCTGCAAGAGGAGTATTGAAAACCTGGTGGACTGTCAGCTTTATTTA  
GCTCACCTAGTGTTTTCAAGAAAATTGAGCCACCGTCTAAGAAATCAAGAGGTTTCACAT  
TAAAATTAGAATTTCTGGCCTCTCTCGATCGGTGAGAATGTGTGGCAATTCTGATCTGCA  
TTTTTCAGAAGAGGACAATCAATTGAACTAAGTAGGGGTTTCTTCTTTTGGCAAGACTTG  
TACTCTCTCACCTGGCCTGTTTCATTTATTTGTATTATCTGCCTGGTCCCTGAGGCGTCT  
GGGTCTCTCCTCTCCCTTGCAAGTTTGGGTTTGAAGCTGAGGAACTACAAAGTTGATGAT  
TTCTTTTTTATCTTTATGCCTGCAATTTTACCTAGCTACCACTAGGTGGATAGTAAATTT  
ATACTTATGTTTCCCTCAAAAAAAAAAAAAA

111/392

## FIGURE 106

METVVIVAIGVLATIFLASFAALVLVCRQRYCRPRDLLQRYDSKPIVDLIGAMETQSEPS  
ELELDDVVITNPHIEAILENEDWIEDASGLMSHCIAILKICHTLTEKLVAMTMGSGAKMK  
TSASVSDIIVVAKRISPRVDDVVKSMYPPLDPKLLDARTTALLSVSHLVLVTRNACHLT  
GGLDWIDQSLSAABEHLEVLREAALASEPDKGLPGPEGFLQEQSAI

112/392

**FIGURE 107**

GCTTCATTTCTCCCGACTCAGCTTCCACCCCTGGGCTTTCCGAGGTGCTTTCGCCGCTGT  
CCCCACCACTGCAGCCATGATCTCCTTAACGGACACGCAGAAAATTGGAATGGGATTAAC  
AGGATTTGGAGTGTTTTTCCTGTTCTTTGGAATGATTCTCTTTTTTGACAAAGCACTACT  
GGCTATTGGAAATGTTTTATTTGTAGCCGGCTTGGCTTTTGTAAATTGGTTTAGAAAGAAC  
ATTCAGATTCTTCTTCCAAAAACATAAAATGAAAGCTACAGGTTTTTTTTCTGGGTGGTGT  
ATTTGTAGTCCTTATTGGTTGGCCTTTGATAGGCATGATCTTCGAAATTTATGGATTTTT  
TCTCTTGTTCAAGGGCTTCTTTCCTGTCTGTTGTTGGCTTTATTAGAAGAGTGCCAGTCCT  
TGGATCCCTCCTAAATTTACCTGGAATTAGATCATTGTAGATAAAGTTGGAGAAAGCAA  
CAATATGGTATAAAACAACAAGTGAATTTGAAGACTCATTTAAAATATTGTGTTATTTATAA  
AGTCATTTGAAGAATATTCAGCACAAAATTAAATTACATGAAATAGCTTGTAATGTTCTT  
TACAGGAGTTTAAAACGTATAGCCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGT  
GAAAACAGGCTTCTACTCAAGTGAAGTAAAGAAGTCAAGCAAGCAAAGTGAAGAGAGGTG  
AAATCCATGTTAATGATGCTTAAGAAACTCTTGAAGGCTATTGTGTTGTTTTCCACAA  
TGTGCGAAACTCAGCCATCCTTAGAGAACTGTGGTGCCTGTTCTTTTCTTTTATTTTG  
AAGGCTCAGGAGCATCCATAGGCATTTGCTTTTTAGAAAGTGTCCACTGCAATGGCAAAAA  
TATTTCCAGTTGCACTGTATCTCTGGAAGTGATGCATGAATTCGATTGGATTGTGTCATT  
TTAAAGTATTAAAACCAAGGAAACCCCAATTTTGATGTATGGATTACTTTTTTTTGNGCN  
CAGGGCC



113/392

## FIGURE 108

MISLTDTQKIGMGLTGFGVFFLFFGMILFFDKALLAIGNVLFVAGLAFVIGLERTFRFFF  
QKHKMKATGFFLGGVFVVLIGWPLIGMIFEIYGFFLLFRGFFPVVVGFIIRVPVLGSLN  
LPGIRSFVDKVGESNNMV

**Important features:**

**Transmembrane domains:**

amino acids 12-30 (typeII), 33-52, 69-89 and 93-109

**N-myristoylation sites:**

amino acids 11-16, 51-56 and 116-121

**Aminoacyl-transfer RNA synthetases class-II protein:**

amino acids 49-59

114/392

**FIGURE 109**

CCAGTCTGTCGCCACCTCACTTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGGTGAGA  
GCACAGAGGAGTGGGCGGGGACCATGCGGGGGACGCGGCTGGCGCTCCTGGCGCTGGTGC  
TGGCTGCCTGCGGAGAGCTGGCGCCGGCCCTGCGCTGCTACGTCTGTCCGGAGCCCACAG  
GAGTGTGCGACTGTGTCACCATCGCCACCTGCACCACCAACGAAACCATGTGCAAGACCA  
CACTCTACTCCCGGGAGATAGTGTACCCCTTCCAGGGGGACTCCACGGTGACCAAGTCCT  
GTGCCAGCAAGTGTAAGCCCTCGGATGTGGATGGCATCGGCCAGACCCTGCCCCTGTCCT  
GCTGCAATACTGAGCTGTGCAATGTAGACGGGGCGCCCGCTCTGAACAGCCTCCACTGCG  
GGGCCCTCACGCTCCTCCCACTCTTGAGCCTCCGACTGTAGAGTCCCCGCCCACCCCAT  
GGCCCTATGCGGCCAGCCCCGAATGCCTTGAAGAAGTGCCCCCTGCACCAGGAAAAAA  
AAAAAAAAA

115/392

## FIGURE 110

MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIV  
YPFQGDSTVTKSCASKCKPSDVGIGQTLFVSCCNTELCNVDGAPALNSLHCGALTLLPL  
LSLRL

**Important features:**

**Signal peptide:**

amino acids 1-17

**N-glycosylation site:**

amino acids 46-49

116/392

## FIGURE 111

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCCGCTTCCTTGAAAAACCCGGCG  
GGCGAGCGAGGCTGCGGGCCGGCCGCTGCCCTTCCCCACACTCCCCGCCGAGAAGCCTCG  
CTCGGCGCCCAACATGGCGGGTGGGCGCTGCGGCCCCGAGCTAACGGCGCTCCTGGCCGC  
CTGGATCGCGGCTGTGGCGGCGACGGCAGGCCCCCGAGGAGGCCGCGCTGCCGCCGGAGCA  
GAGCCGGGTCCAGCCCATGACCGCCTCCAACCTGGACGCTGGTGATGGAGGGCGAGTGGAT  
GCTGAAATTTTACGCCCCATGGTGTCCATCCTGCCAGCAGACTGATTCAGAATGGGAGGC  
TTTTGCAAAGAATGGTGAAATACTTCAGATCAGTGTGGGAAGGTAGATGTCATTCAAGA  
ACCAGGTTTGAGTGGCCGCTTCTTTGTCAACACTCTCCAGCATTTTTTCATGCAAAGGA  
TGGGATATTCCGCCGTTATCGTGGCCCAGGAATCTTCGAAGACCTGCAGAATTATATCTT  
AGAGAAGAAATGGCAATCAGTCGAGCCTCTGACTGGCTGGAAATCCCCAGCTTCTCTAAC  
GATGTCTGGAATGGCTGGTCTTTTAGCATCTCTGGCAAGATATGGCATCTTCACAACTA  
TTTCACAGTGACTCTTGGAAATTCCTGCTTGGTGTTCTTATGTGTTTTTCGTCATAGCCAC  
CTTGGTTTTTGGCCTTTTTATGGGTCTGGTCTTGGTGTTAATATCAGAATGTTTTCTATGT  
GCCACTTCCAAGGCATTTATCTGAGCGTTCTGAGCAGAATCGGAGATCAGAGGAGGCTCA  
TAGAGCTGAACAGTTGCAGGATGCGGAGGAGGAAAAAGATGATTCAAATGAAGAAGAAAA  
CAAAGACAGCCTTGTAGATGATGAAGAAGAGAAAGAAAGATCTTGGCGATGAGGATGAAGC  
AGAGGAAGAAGAGGAGGAGGACAACCTTGGCTGCTGGTGTGGATGAGGAGAGAAGTGAGGC  
CAATGATCAGGGGCCCCCAGGAGAGGACGGTGTGACCCGGGAGGAAGTAGAGCCTGAGGA  
GGCTGAAGAAGGCATCTCTGAGCAACCCTGCCAGCTGACACAGAGGTGGTGGAAAGACTC  
CTTGAGGCAGCGTAAAGTCAGCATGCTGACAAGGGACTGTAGATTTAATGATGCGTTTT  
CAAGAATACACACCAAACAATATGTCAGCTTCCCTTTGGCCTGCAGTTTGTACCAAATC  
CTTAATTTTTCTGAATGAGCAAGCTTCTCTTAAAGATGCTCTCTAGTCATTTGGTCTC  
ATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGTGACAATCAGGATATAG  
AAAAACAAACGTAGTGTTGGGATCTGTTTGGGAGACTGGGATGGGAACAAGTTCATTTACT  
TAGGGGTGAGAGAGTCTCGACCAGAGGAGGCCATTCCCAGTCCTAATCAGCACCTTCCAG  
AGACAAGGCTGCAGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTTGGCTCCTGAGCATC  
CCCAAAGTGTAACGTAGAAGCCTTGCAATCCTTTTTCTTGTGTAAAGTATTTATTTTTGTCA  
AATTGCAGGAAACATCAGGCACCACAGTGCATGAAAAATCTTTCACAGCTAGAAATTGAA  
AGGGCCTTGGGTATAGAGAGCAGCTCAGAAGTCATCCAGCCCTCTGAATCTCCTGTGCT  
ATGTTTTATTTCTTACCTTTAATTTTTCCAGCATTTCCACCATGGGCATTCAGGCTCTCC  
ACACTCTTCACTATTATCTCTTGGTCAGAGGACTCCAATAACAGCCAGGTTTACATGAAC  
TGTGTTTGTTCATTCTGACCTAAGGGGTTAGATAATCAGTAACCATAACCCCTGAAGCT  
GTGACTGCCAAACATCTCAAATGAAATGTTGTGGCCATCAGAGACTCAAAGGAAGTAAG  
GATTTTACAAGACAGATTAAAAAAAATTTGTTTTGTCCAAAATATAGTTGTTGTTGATTT  
TTTTTTAAGTTTTCTAAGCAATATTTTTCAAGCCAGAAGTCCTCTAAGTCTTGCCAGTAC  
AAGGTAGTCTTGTGAAGAAAAGTTGAATACTGTTTTGTTTTCATCTCAAGGGTTCCCTG  
GGTCTTGAACACTTTTAATAATAACTAAAAAACCACTTCTGATTTTCCTTCAGTGATGTG  
CTTTTGGTGAAAGAATTAAATGAACCTCAGTACCTGAAAGTGAAAGATTTGATTTTGTTC  
CATCTTCTGTAATCTTCCAAGAATTATATCTTTGTAAATCTCTCAATACTCAATCTACT  
GTAAGTACCCAGGAGGCTAATTTCTTT

117/392

**FIGURE 112**

MAGGRCGPQLTALLAAWIAAVAATAGPEEAALPPEQSRVQPM TASNWTLVMEGEWMLKFY  
APWCPSCQOTDSEWEAFAKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGIFR  
RYRGP G IFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVT  
LGIPAWCSYVFFVIATLVFGLFMGLVLVVISECFYVPLPRHLSERSEQNRRSEEAHRAEQ  
LQDAEEEEKDDSN EEEENKDSLVDDEEEKEDLGDEDEAE EEEEEEDNLAAGVDEERSEANDQG  
PPGEDGVTR EEEVEPEEAEEGISEQPCPADTEVVEDSLRQRKSQHADKGL

**Important features:**

**Signal peptide:**

amino acids 1-22

**Transmembrane domain:**

amino acids 191-211

**N-glycosylation site:**

amino acids 46-49

**Thioredoxin family proteins:** (homologous region to disulfide isomerase)

amino acids 56-72

**Flavodoxin proteins:**

amino acids 173-187

118/392

## FIGURE 113

GAGGAACCTACCGGTACCGGCCGCGCGCTGGTAGTCGCCGGTGTGGCTGCACCTCACCAA  
TCCCGTGCGCCGCGGCTGGGCGGTGCGAGAGTGCCTGTGCTTCTCTCCTGCACGCGGTGC  
TTGGGCTCGGCCAGGCGGGGTCCGCCGCCAGGGTTTGAGGATGGGGGAGTAGCTACAGGA  
AGCGACCCCGCGATGGCAAGGTATATTTTTGTGGAATGAAAAGGAAGTATTAGAAATGAG  
CTGAAGACCATTACAGATTAATATTTTTGGGGACAGATTTGTGATGCTTGATTACCCCT  
TGAAGTAATGTAGACAGAAGTTCTCAAATTTGCATATTACATCAACTGGAACCAGCAGTG  
AATCTTAATGTTCACTTAAATCAGAACTTGCATAAGAAAAGAGAATGGGAGTCTGGTTAAA  
TAAAGATGACTATATCAGAGACTTGAAAAGGATCATTCTCTGTTTTCTGATAGTGTATAT  
GGCCATTTTAGTGGGCACAGATCAGGATTTTTACAGTTTACTTGGAGTGTCCAAAACCTGC  
AAGCAGTAGAGAAATAAGACAAGCTTTCAAGAAATTGGCATTGAAGTTACATCCTGATAA  
AAACCCGAATAACCCAAATGCACATGGCGATTTTTTAAAAATAAATAGAGCATATGAAGT  
ACTCAAAGATGAAGATCTACGGAAAAAGTATGACAAATATGGAGAAAAGGGACTTGAGGA  
TAATCAAGGTGGCCAGTATGAAAGCTGGAACATTTATCGTTATGATTTTGGTATTTATGA  
TGATGATCCTGAAATCATAACATTGGAAAGAAGAGAATTTGATGCTGCTGTTAATTCTGG  
AGAACTGTGGTTTGTAAATTTTTACTCCCCAGGCTGTTCCACTGCCATGATTTAGCTCC  
CACATGGAGAGACTTTGCTAAAGAAGTGGATGGGTACTTCGAATTGGAGCTGTTAACTG  
TGGTGATGATAGAATGCTTTGCCGAATGAAAGGAGTCAACAGCTATCCCAGTCTCTTCAT  
TTTTCGGTCTGGAATGGCCCCAGTGAAATATCATGGAGACAGATCAAAGGAGAGTTTAGT  
GAGTTTGCATGCAGCATGTTAGAAGTACAGTGACAGAACTTTGGACAGGAAATTTTGT  
CAACTCCATACAACTGCTTTTGCTGCTGGTATTGGCTGGCTGATCACTTTTTGTTCAA  
AGGAGGAGATTGTTTGACTTCACAGACAGCACTCAGGCTTAGTGGCATGTTGTTTCTCAA  
CTCTTGGATGCTAAAGAAATATATTGGGAAGTAATACATAATCTTCCAGATTTTGAAC  
ACTTTTCGGCAACACACTAGAGGATCGTTTGGCTCATCATCGGTGGCTGTTATTTTTTCA  
TTTTTGAAAAAATGAAAATTCAAATGATCCTGAGCTGAAAAAACTAAAACTCTACTTAA  
AAATGATCATATTCAAGTTGGCAGGTTTGACTGTTTCTCTGCACCAGACATCTGTAGTAA  
TCTGTATGTTTTTTCAGCCGTCTCTAGCAGTATTTAAAGGACAAGGAACCAAAGAATATGA  
AATTCATCATGGAAGAAGATTCTATATGATATACTTGCCTTTGCCAAAGAAGTGTGAA  
TTCTCATGTTACCACGCTTGGACCTCAAATTTTCTGCCAATGACAAAGAACCATGGCT  
TGTTGATTTCTTTGCCCCCTGGTGTCCACCATGTGAGCTTTACTACCAGAGTTACGAAG  
AGCATCAAATCTTCTTATGGTCAGCTTAAGTTTGGTACACTAGATTGTACAGTTTCATGA  
GGGACTCTGTAACATGTATAACATTAGGCTTATCCAACAACAGTGGTATTCAACCAGTC  
CAACATTCATGAGTATGAAGGACATCACTCTGCTGAACAAATCTTGGAGTTCATAGAGGA  
TCTTATGAATCCTTCAGTGGTCTCCCTTACACCCACCACCTTCAACGAACTAGTTACACA  
AAGAAAACACAACGAAGTCTGGATGGTTGATTTCTATTCTCCGTGGTGTATCCTTGCCA  
AGTCTTAATGCCAGAATGGAAAAGAATGGCCCGGACATTAACCTGGACTGATCAACGTGGG  
CAGTATAGATTGCCAACAGTATCATTCTTTTTGTGCCCGAGGAAAACGTTCAAAGATACCC  
TGAGATAAGATTTTTTCCCCCAAATCAAATAAGCTTATCAGTATCACAGTTACAATGG  
TTGGAATAGGGATGCTTATCCCTGAGAATCTGGGGTCTAGGATTTTTACCTCAAGTATC  
CACAGATCTAACACCTCAGACTTTTCACTGAAAAAGTTCTACAAGGGAAAAATCATTGGGT  
GATTGATTTCTATGCTCCTTGGTGTGGACCTTGCCAGAATTTTGTCTCAGAATTTGAGCT  
CTTGGCTAGGATGATTAAAGGAAAAGTGAAAGCTGGAAAAGTAGACTGTGAGGCTTATGC  
TCAGACATGCCAGAAAGCTGGGATCAGGGCCTATCCAACCTGTTAAGTTTATTTCTACGA  
AAGAGCAAAGAGAAATTTTCAAGAAGAGCAGATAAATAACCAGAGATGCAAAAGCAATCGC  
TGCCTTAATAAGTGAAAAATTTGGAACTCTCCGAAATCAAGGCAAGAGGAATAAGGATGA  
ACTTTGATAATGTTGAAGATGAAGAAAAAGTTTAAAAGAAATTTCTGACAGATGACATCAG  
AAGACACCTATTTAGAATGTTACATTTATGATGGGAATGAATGAACATTATCTTAGACTT

119/392

GCAGTTGTACTGCCAGAATTATCTACAGCACTGGTGTAAAAGAAGGGTCTGCAAACCTTTT  
TCTGTAAAGGGCCGGTTTATAAATATTTTAGACTTTGCAGGCTATAATATATGGTTCACA  
CATGAGAACAAGAATAGAGTCATCATGTATTCTTTGTTATTTGCTTTTAAACAACCTTTAA  
AAAATATTTAAACGATTCTTAGCTCAGAGCCATACAAAAGTAGGCTGGATTCCAGTCCATG  
GACCATAGATTGCTGTCCCCCTCGACGGACTTATAATGTTTCAGGTGGCTGGCTTGAACA  
TGAGTCTGCTGTGCTATCTACATAAATGTCTAAGTTGTATAAAGTCCACTTCCCTTCAC  
GTTTTTTGGCTGACCTGAAAAGAGGTAACCTTAGTTTTTTGGTCACTTGTTCTCCTAAAAAT  
GCTATCCCTAACCATATATTTATATTTTCGTTTTAAAAACACCCATGATGTGGCACAGTAA  
ACAAACCCTGTTATGCTGTATTATTATGAGGAGATTCTTCATTGTTTTCTTTCCTTCTCA  
AAGGTTGAAAAAATGCTTTTAATTTTTTACAGCCGAGAAACAGTGCAGCAGTATATGTGC  
ACACAGTAAGTACACAAATTTGAGCAACAGTAAGTGCACAAATTCGTAGTTTGCTGTAT  
CATCCAGGAAAACCTGAGGGAAAAAATTATAGCAATTAAGTGGGCATTGTAGAGTATCC  
TAAATATGTTATCAAGTATTTAGAGTTCTATATTTTAAAGATATATGTGTTTCATGTATTT  
TCTGAAATTGCTTTCATAGAAATTTTCCCACTGATAGTTGATTTTTGAGGCATCTAATAT  
TTACATATTTGCCTTCTGAACCTTGTTTTGACCTGTATCCTTTATTTACATTGGGTTTTT  
CTTTCATAGTTTTGGTTTTTCACTCCTGTCCAGTCTATTTATTATTCAAATAGGAAAAAT  
TACTTTACAGGTTGTTTTACTGTAGCTTATAATGATACTGTAGTTATTCCAGTTACTAGT  
TTACTGTCAGAGGGCTGCCTTTTTTCAGATAAATATTGACATAATAACTGAAGTATTTTTT  
ATAAGAAAATCAAGTATATAAATCTAGGAAAGGGATCTTCTAGTTTCTGTGTTGTTTAGA  
CTCAAAGAATCACAAATTTGTCAGTAACATGTAGTTGTTTAGTTATAAATTCAGAGTGTAC  
AGAATGGTAAAAATTCCAATCAGTCAAAGAGGTCAATGAATTAAAGGCTTGCAACTTT  
TTCAAAAAAAAAAAAAAAAAA

120/392

**FIGURE 114**

MGVWLNKDDYIRD LKRIILCFLIVYMAILVGT DQDFYSLLGVSKTASSREIRQAFKKLAL  
KLHPDKNPNNPNAHGDF LKINRAYEVLKDEDLRKKYDKYGEK GLEDNQGGQYESWNYRY  
DFGIYDDDPEIITLERREFDAAVNSGELW FVNFYSPGCSHCHDLAPTWRDFAKEVDGLLR  
IGAVNCGDDRMLCRMKG VNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTEL  
WTGNFVNSIQTAFAAGIGWLITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHN  
LPDFELLSANTLEDRLA HHRWLLFFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSA  
PDICSNLYVFQPSLAVF KGQGTKEYEIIHHGKKILYDILAFAKESVNSHVTTLGPNFPAN  
DKEPWLVDFFAPWCP PCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTT  
VVFNQSNIEHEGHHSAEQILEFIEDLMNPSV VSLTPTTFNELVTQRKHNEVWMVDFYSP  
WCHPCQVLMPEWKRMARTLTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQ  
YHSYNGWNRDAYSLRIWGLGFLPQVSTD LTPQTFSEKVLQGKNHWVIDFYAPWCGPCQNF  
APEFELLARMIKGKVKAGKVDCQAYAQTCQKAGIRAYPTVKFYFYERAKRNFQEEQINTR  
DAKAIAALISEKLETLRNQGKR NKDEL

**Important features:**

**Endoplasmic reticulum targeting sequence:**

amino acids 744-747

**Cytochrome c family heme-binding site signature:**

amino acids 158-163

**Nt-dnaJ domain signature:**

amino acids 77-96

**N-glycosylation site:**

amino acids 484-487



121/392

## FIGURE 115

GCGGGCTGTTGACGGCGCTGCGATGGCTGCCTGCGAGGGCAGGAGAAGCGGAGCTCTCGG  
TTCCTCTCAGTCGGACTTCCTGACGCCGCCAGTGGGCGGGGCCCCCTTGGGCGCTCGCCAC  
CACTGTAGTCATGTACCCACCGCCGCCGCCGCCCTCATCGGGACTTCATCTCGGTGAC  
GCTGAGCTTTGGCGAGAGCTATGACAACAGCAAGAGTTGGCGGCGGCGCTCGTGCTGGAG  
GAAATGGAAGCAACTGTCGAGATTGCAGCGGAATATGATTCTCTTCTCTTGCCTTTCT  
GCTTTTCTGTGGACTCCTCTTCTACATCAACTTGGCTGACCATTGGAAGCTCTGGCTTT  
CAGGCTAGAGGAAGAGCAGAAGATGAGGCCAGAAATTGCTGGGTAAACCAGCAAATCC  
ACCCGTCTTACCAGCTCCTCAGAAGGCGGACCCGACCCTGAGAACTTACCTGAGATTTT  
GTCACAGAAGACACAAAGACACATCCAGCGGGGACCACCTCACCTGCAGATTAGACCCCC  
AAGCCAAGACCTGAAGGATGGGACCCAGGAGGAGGCCACAAAAGGCAAGAAGCCCCGTGT  
GGATCCCCCGCCGGAAGGAGATCCGCAGAGGACAGTCATCAGCTGGAGGGGAGCGGTGAT  
CGAGCCTGAGCAGGGCACCGAGCTCCCTTCAAGAAGAGCAGAAAGTGCACCAAGCCTCC  
CCTGCCACCGGCCAGGACACAGGGCACACCAGTGTCATCTGAACTATCGCCAGAAGGGCGT  
GATTGACGTCTTCTGTCATGCATGGAAGGATACCGCAAGTTTGCATGGGGCCATGACGA  
GCTGAAGCCTGTGTCCAGGTCCTTCAGTGAGTGGTTTGGCCTCGGTCTCACACTGATCGA  
CGCGCTGGACACCATGTGGATCTTGGGTCTGAGGAAAGAATTTGAGGAAGCCAGGAAGTG  
GGTGTGAAGAAGTTACACTTTGAAAAGGACGTGGACGTCAACCTGTTTGAGAGCACGAT  
CCGCATCCTGGGGGGGCTCCTGAGTGCCTACCACCTGTCTGGGGACAGCCTCTTCTGAG  
GAAAGCTGAGGATTTTGGAAATCGGCTAATGCCTGCCTTCAGAACACCATCCAAGATTCC  
TACTCGGATGTGAACATCGGTACTGGAGTTGCCACCCGCCACGGTGGACCTCCGACAG  
CACTGTGGCCGAGGTGACCAGCATTGAGCTGGAGTTCCGGGAGCTCTCCCGTCTCACAGG  
GGATAAGAAGTTTCAGGAGGCAGTGGAGAAGGTGACACAGCACATCCACGGCCTGTCTGG  
GAAGAAGGATGGGCTGGTGCCCATGTTTCATCAATACCCACAGTGGCCTCTTACCCACCT  
GGGCGTATTACGCTGGGCGCCAGGGCCGACAGCTACTATGAGTACCTGCTGAAGCAGTG  
GATCCAGGGCGGGAAGCAGGAGACACAGCTGCTGGAAGACTACGTGGAAGCCATCGAGGG  
TGTCAGAAGCACCCTGCTGCGGCACTCCGAGCCCAGTAAGCTCACCTTTGTGGGGGAGCT  
TGCCACCGGCCGCTTCAGTGCCAAGATGGACCACCTGGTGTGCTTCTGCCAGGGACGCT  
GGCTCTGGGCGCTTACCACGGCCTGCCCGCCAGCCACATGGAGCTGGCCCAGGAGCTCAT  
GGAGACTTGTACCAGATGAACCGGCAGATGGAGACGGGGCTGAGTCCCGAGATCGTGCA  
CTTCAACCTTTACCCCCAGCCGGGCCGTCGGGACGTGGAGGTCAAGCCAGCAGACAGGCA  
CAACCTGCTGCGGCCAGAGACCGTGGAGAGCCTGTTCTACCTGTACCGCTCACAGGGGA  
CCGCAAATACCAGGACTGGGGCTGGGAGATTCTGCAGAGCTTCAGCCGATTACACGGGT  
CCCCTCGGGTGGCTATTCTTCCATCAACAATGTCCAGGATCCTCAGAAGCCCCAGCCTAG  
GGACAAGATGGAGAGCTTCTTCTGGGGGAGACGCTCAAGTATCTGTTCTTGCTCTTCTC  
CGATGACCCAAACCTGCTCAGCCTGGACGCCTACGTGTTCAACACCGAAGCCACCTCT  
GCCTATCTGGACCCCTGCCTAGGGTGGATGGCTGCTGGTGTGGGGACTTCGGGTGGGCAG  
AGGCACCTTGCTGGGTCTGTGGCATTTCCTCAAGGGCCCACGTAGCACCGGCAACCGCCAA  
GTGGCCCAGGCTCTGAACTGGCTCTGGGCTCCTCCTCGTCTCTGCTTTAATCAGGACACC  
GTGAGGACAAGTGAGGCCGTGAGTCTTGGTGTGATGCGGGGTGGGCTGGGCCGCTGGAGC  
CTCCGCTGCTTCTTCCAGAAGACACGAATCATGACTCACGATTGCTGAAGCCTGAGCAG  
GTCTCTGTGGGCCGACCAGAGGGGGGCTTCGAGGTGGTCCCTGGTACTGGGGTGACCGAG  
TGGACAGCCCAGGGTGACGCTCTGCCCCGGGCTCGTGAAGCCTCAGATGTCCCAATCCAA  
GGGTCTGGAGGGGCTGCCGTGACTCCAGAGGCCTGAGGCTCCAGGGCTGGCTCTGGTGT  
TACAAGCTGGACTCAGGGATCCTCCTGGCCGCCCCGAGGGGGCTTGGAGGGCTGGACGG  
CAAGTCCGTCTAGCTCACGGGCCCCCTCAGTGGAATGGGTCTTTTCGGTGGAGATAAAAG  
TTGATTTGCTCTAACCGCAA

122/392

**FIGURE 116**

MAACEGRRSGALGSSQSDFLTPPVGGAPWAVATTVMYPPPPPPPHRDFISVTLSFGESY  
DNSKSWRRRSCWRKWKQLSRLQRMILFLLAFLLFCGLLFYINLADHWKALAFRLEEEQK  
MRPEIAGLKPANPPVLPAPQKADTDPENLPEISSQKTQRHIQRGPPHLQIRPPSQDLKDG  
TQEEATKRQEAPVDPRPEGDPQRTVISWRGAVIEPEQGTELPSRRAEVPTKPPLPPARTQ  
GTPVHLNRYQKGVIDVFLHAWKGYRKFAWGHDELKPVSRSFSEWFGLGLTLIDALDTMWI  
LGLRKEFEEARKWVSKLHFEKDVDVNLFESTITRILGGLLSAYHLSGDSLFLRKAEDFGN  
RLMPAFRTPSKIPYSDVNIGTGVAHPPRWTSdstVAEVTsiQLEFRELSRLTGDKKFQEA  
VEKVTQHIHGLSGKKDGLVPMFINTHSGLFTHLGVFTLGARADSYEYLLKQWIQGGKQE  
TQLLEDYVEAIEGVRTLLRHSEPSKLTfVgELAHGRfSAKMDHLVCFLPGTLALGVYHG  
LPASHMELAQELMETCYQMNRQMETGLSPEIVHFNLYPQGRRDVEVKPADRHNLRLRPET  
VESLFYLYRVTGDRKYQDWGWEILQSFSRfTRVPSGGYSSINNVDpQKPEPRDKMESFF  
LGETLKYLfLLfSDDPNLLSLDAYVFNTeAHPLPIWTPA

**Important features of the protein:**

**Transmembrane domain:**

amino acids 21-40 and 84-105 (type II)

123/392

**FIGURE 117**

GTGGGATTTATTTGAGTGCAAGATCGTTTTCTCAGTGGTGGTGGAAGTTGCCTCATCGCA  
GGCAGATGTTGGGGCTTTGTCCGAACAGCTCCCCCTCTGCCAGCTTCTGTAGATAAGGGTT  
AAAACTAATATTTATATGACAGAAGAAAAAGATGTCATTCCGTAAAGTAAACATCATCA  
TCTTGGTCCTGGCTGTTGCTCTCTTCTTACTGGTTTTGACCATAACTTCCTCAGCTTGA  
GCAGTTTGTTAAGGAATGAGGTTACAGATTGAGGAATTGTAGGGCCTCAACCTATAGACT  
TTGTCCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGAGATTCTGTGGTCA  
TCGCTGCATCTGAAGACAGGCTTGGGGGGGCCATTGCAGCTATAAACAGCATTACAGCACA  
ACACTCGCTCCAATGTGATTTTCTACATTGTTACTCTCAACAATACAGCAGACCATCTCC  
GGTCTGGCTCAACAGTGATTCCCTGAAAAGCATCAGATACAAAATTGTCAATTTTGACC  
CTAACTTTTGAAGGAAAAGTAAAGGAGGATCCTGACCAGGGGAATCCATGAAACCTT  
TAACCTTTGCAAGGTTCTACTTGCCAATTCTGGTTCCCAGCGCAAAGAAGGCCATATACA  
TGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTACAATACAGCACTGAAGC  
CAGGACATGCAGCTGCATTTTTCAGAAGATTGTGATTGAGCCTCTACTAAAGTTGTCATCC  
GTGGAGCAGGAAACCAGTACAATTACATTGGCTATCTTGACTATAAAAAGGAAAGAATTC  
GTAAGCTTTCCATGAAAGCCAGCACTTGCTCATTTAATCCTGGAGTTTTTGTGCAAACC  
TGACGGAATGGAAACGACAGAATATACTAACCAACTGGAAAAATGGATGAAACTCAATG  
TAGAAGAGGGACTGTATAGCAGAACCCTGGCTGGTAGCATCACAACACCTCCTCTGCTTA  
TCGTATTTTATCAACAGCACTCTACCATCGATCCTATGTGGAATGTCGGCCACCTTGGTT  
CCAGTGCTGGAAAACGATATTACCTCAGTTTGTAAAGGCTGCCAAGTTACTCCATTGGA  
ATGGACATTTGAAGCCATGGGGAAGGACTGCTTCATATACTGATGTTTGGGAAAAATGGT  
ATATTCCAGACCCAACAGGCAAATTC AACCTAATCCGAAGATATACCGAGATCTCAAACA  
TAAAGTGAAACAGAATTTGAACTGTAAGCAAGCATTCTCAGGAAGTCCTGGAAGATAGC  
ATGCATGGGAAGTAACAGTTGCTAGGCTTCAATGCCTATCGGTAGCAAGCCATGGAAAAA  
GATGTGTCAGCTAGGTAAAGATGACAACTGCCCTGTCTGGCAGTCAGCTTCCCAGACAG  
ACTATAGACTATAAATATGTCTCCATCTGCCTTACCAAGTGTCTTACTACAATGCTG  
AATGACTGGAAAGAAGAACTGATATGGCTAGTTGAGCTAGCTGGTACAGATAATTCAAAA  
CTGCTGTTGGTTTTAATTTTGTAACTGTGGCCTGATCTGTAAATAAACTTACATTTTTC

124/392

**FIGURE 118**

MSFRKVNIIILVLAVALFLLVLHNFSLSLSLLRNEVTDSGIVGPQPIDFVPNALRHAVD  
GRQEEIPVVIAASEDRLGGAIAAINSIOHNTRSNVIFYIVTLNNTADHLRSWLNSDSLKS  
IRYKIVNFDPKLLEGKVKEPDQGESMKPLTFARFYLPILVPSAKKAIYMDDDVIVQGDI  
LALYNTALKPGHAAAFSEDCDSASTKVVIRGAGNQYNYIGYLDYKKERIRKLSMKASTCS  
FNPGVFVANLTEWKRQNTITNQLEKWMKLNVEEGLYSRTLGSITTPPLLIVFYQQHSTID  
PMWNVRLGSSAGKRYSPQFVKAACKLLHWNGHLKPWGRTASYTDVWEKWYIPDPTGKFNL  
IRRYTEISNIK

125/392

**FIGURE 119**

CCATCCCTGAGATCTTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATACCAG  
ATCTCACCAGAGAGTCGCAGACACTATGCTGCCTCCCATGGCCCTGCCCAGTGTGTCTCTG  
GATGCTGCTTTCCTGCCTCATTCTCCTGTGTCTCAGGTTCAAGGTGAAGAAACCCAGAAGGA  
ACTGCCCTCTCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTA  
TGCCTTGTTTTTGTACCAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCC  
CTCTGGAAACTGGTGTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCTCCCTGGT  
GAGGAGCATTAGTAACAGCTACTCATACTCTGGATTGGGCTCCATGACCCACACAGGG  
CTCTGAGCCTGATGGAGATGGATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGC  
ATGGGAGAAAAATCCCTCCACCATCTTAAACCCTGGCCACTGTGGGAGCCTGTCAAGAAG  
CACAGGATTTCTGAAGTGGAAGATTATACTGTGATGCAAAGTTACCCTATGTCTGCAA  
GTTCAAGGACTAGGGCAGGTGGGAAGTCAGCAGCCTCAGCTTGGCGTGCAGCTCATCATG  
GACATGAGACCAGTGTGAAGACTCACCTGGAAGAGAATATTCTCCCCAACTGCCCTAC  
CTGACTACCTTGTATGATCCTCCTTCTTTTTCCTTTTTCTTCACCTTCATTTCAGGCTT  
TTCTCTGTCTTCCATGTCTTGAGATCTCAGAGAATAATAATAAAAATGTTACTTTATAAA  
AAAAAAAAAAAAAAAAAAAA

126/392

## FIGURE 120

MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKS  
WMDADLACQKRPSGKLVSVLSGAEGSFVSSLVRSISNSYSYIWIGLHDPTQGSEPDGDGW  
EWSSTDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWKDYNCDAKLPYVCKFKD

Important features:

Signal peptide:

amino acids 1-26

C-type lectin domain signature:

amino acids 146-171

127/392

**FIGURE 121**

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTT  
TGGGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGCTCTAG  
AACAAATTCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAG  
ATGGCTGAGATGGACAGAATGCTTTATTTTGGAAAGAAAATGTTCTAGGTCAAACCTGA  
GTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAAATCTGGACAAGTCTTTTCATG  
TGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCC  
CCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTG  
ATCGCGCCTGGAGAAACAGTGTACTATTCTGTGCAATACCAGGGGGAGTACGAGAGCCTG  
TACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCTTGAGTGT  
GATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTCGTGTCAGGGCCACATTG  
GGCTCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATC  
CTTACCCGACCTGGGATGGAGATCACCAGAGATGGCTTCCACCTGGTTATTGAGCTGGAG  
GACCTGGGGCCCCAGTTTGAGTTCCTTGTCCTTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAG  
GAACATGTCAAAATGGTGAGGAGTGGGGGTATTCAGTGCACCTAGAAACCATGGAGCCA  
GGGGCTGCATACTGTGTGAAGGCCAGACATTCGTGAAGGCCATTGGGAGGTACAGCGCC  
TTCAGCCAGACAGAATGTGTGGAGGTGCAAGGAGAGGCCATTCCCTGGTACTGGCCCTG  
TTTGCCTTTGTTGGCTTCATGCTGATCCTTGTGGTTCGTGCCACTGTTTCGTCTGAAAAATG  
GGCCGGCTGCTCCAGTACTCCTGTTGCCCGTGGTGGTTCCTCCAGACACCTTGAAAATA  
ACCAATTCACCCAGAGTTAATCAGCTGCAGAGGGAGGAGGTGGATGCCTGTGCCACG  
GCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGATCTCATAAGTTTGCAGGAGGGCC  
CAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACCATGAGGGGACAAGTTGTGTT  
TCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGAGCCTGTTGTCTACAAGTC  
TAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACTGACTGAGGCTTAGGG  
GATGTGACCTCTAGACTGGGGCTGCCACTTGCTGGCTGAGCAACCCTGGGAAAAGTGAC  
TTCATCCCTTCGGTCCCTAAGTTTTCTCATCTGTAATGGGGGAATTACCTACACACCTGCT  
AAACACACACACAGAGTCTCTCTCTATATATACACACGTACACATAAATACACCCAGC  
ACTTGCAAGGCTAGAGGGAACTGGTGACACTCTACAGTCTGACTGATTAGTGTCTCTG  
GAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGTGGCTTG  
GAGAGCCCACTTTCCAGAAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGGTGTT  
GAGTTCACTTCAAGCCCAATGCCGGTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGTAG  
GTGACCTGGAGGAAGGTCACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC  
CATGAACTACTGTAAAGTGTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTAT  
GTGTGCAATGCGACGAGAATGCAGAAGTCAGTAAATGTGCATGTTTGTGTGCTCCTTT  
TTTCTGTTGGTAAAGTACAGAATTCAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGT  
AAAAAAAAAAAAAAAA

128/392

**FIGURE 122**

MQTFTMVLEEIIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLMLWSPVIAP  
GETVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECVTDITATVPYNLRVRATLGSQ  
TSAWSILKHPFNRNSTILTRPGMEITKDGFLVIELEDLGPQFEFLVAYWRREPGAEHV  
KMVRSGGIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPVLALFAF  
VGFMLILVVVPLFVWKMGRLLQYSCCPVVVLPTLKITNSPQKLISCRREEVDACATAVM  
SPEELLRAWIS

**Important features:**

**Signal peptide:**

amino acids 1-29

**Transmembrane domain:**

amino acids 230-255

**N-glycosylation sites:**

amino acids 40-43 and 134-137

**Tissue factor proteins homology:**

amino acids 92-119

**Integrins alpha chain protein homology:**

amino acids 232-262



129/392

**FIGURE 123**

CGGACGCGTGGGCCGCCACCTCCGGAACAAGCCATGGTGGCGGCGACGGTGGCAGCGGCCG  
TGGCTGCTCCTGTGGGCTGCGGCCTGCGCGCAGCAGGAGCAGGACTTCTACGACTTCAAG  
GCGGTCAACATCCGGGGCAAACCTGGTGTGCTGCGTGGAGAAGTACCGCGGATCGGTGTCCCTG  
GTGGTGAATGTGGCCAGCGAGTGGGCTTCACAGACCAGCACTACCGAGCCCTGCAGCAG  
CTGCAGCGAGACCTGGGCCCCCACCCTTTAACGTGCTCGCCTTCCCCTGCAACCAGTTT  
GGCCAACAGGAGCCTGACAGCAACAAGGAGATTGAGAGCTTTGCCCCCGGCACCTACAGT  
GTCTCATTTCCCATGTTTAGCAAGATTGCAGTCACCGGTACTGGTGGCCATCCTGCCTTC  
AAGTACCTGGCCCAGACTTCTGGGAAGGAGCCCACCTGGAACCTTCTGGAAGTACCTAGTA  
GCCCCAGATGGAAAGGTGGTAGGGGCTTGGGACCCAACTGTGTGAGTGGAGGAGGTCAGA  
CCCCAGATCACAGCGCTCGTGAGGAAGCTCATCCTACTGAAGCGAGAAGACTTATAACCA  
CCGCGTCTCCTCCTCCACCACCTCATCCCGCCCACCTGTGTGGGGCTGACCAATGCAAAC  
TCAAATGGTGCTTCAAAGGGAGAGACCCACTGACTCTCCTTCCTTTACTCTTATGCCATT  
GGTCCCATCATTTCTTGTGGGGGAAAAATTCTAGTATTTTGATTATTTGAATCTTACAGCA  
ACAAATAGGAACTCCTGGCCAATGAGAGCTCTTGACCAGTGAATCACCAGCCGATACGAA  
CGTCTTGCCAACAAAAATGTGTGGCAAATAGAAGTATATCAAGCAATAATCTCCCACCCA  
AGGCTTCTGTAAACTGGGACCAATGATTACCTCATAGGGCTGTTGTGAGGATTAGGATGA  
AATACCTGTGAAAGTGCCTAGGCAGTGCCAGCCAAATAGGAGGCATTCAATGAACATTTT  
TTGCATATAAACCAAAAAATAACTTGTTATCAATAAAAACTTGCATCCAACATGAATTTCT  
CAGCCGATGATAATCCAGGCCAAAGGTTTAGTTGTTGTTATTTCTCTGTATTATTTCT  
TCATTACAAAAGAAATGCAAGTTCATTGTAACAATCCAAACAATACCTCACGATATAAAA  
TAAAAATGAAAGTATCCTCCTCAAAA

130/392

## FIGURE 124

MVAATVAAAWLLLWAAACAQQEQDFYDFKAVNIRGKLVSLEKYRGSVSLVVNVASECGFT  
DQHYRALQQLQRDLGPHHFNVLAFFPCNQFGQQEPDSNKEIESFARRTYSVSFPMFSKIAV  
TGTGAHPAFKYLAQTSGKEPTWNFWKYLVPDGKVVGAWDPTVSVEEVRPQITALVRKLI  
LLKREDL

131/392

**FIGURE 125**

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTGGGAGGGGGCAGGATGGGAGGG  
AAAGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAG  
AAACCGATCAGGCATGGAACTCCCCCTTCGTCACCTGTTCTTGCCCCCTGGTGTTCCT  
GACAGGTCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCTATTCCCAGGGCC  
ACCAGAAGCTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGACAGCGATGGAT  
GCTGGTGGGCGCCCCCTGGGATGGGCCTTCAGGCGACCGAGGGGGGACGTTTATCGCTG  
CCCTGTAGGGGGGGGCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTACCAACT  
GGGAAATTCATCTCATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGACAGA  
TGGTGATGGGGGATTCATGGTGAGCTAAGGAGAGGGTGGTGGCAGTGTCTCTGAAGGTCC  
ATAAAAGAAAAAGAGAAGTGTGGTAAGGGAAAATGGTCTGTGTGGAGGGGTCAAGGAGT  
TAAAAACCCTAGAAAGCAAAAGGTAGGTAATGTCAGGGAGTAGTCTTCATGCCTCCTTCA  
ACTGGGAGCATGTTCTGAGGGTGCCCTCCCAAGCCTGGGAGTAACTATTTCCCCCATCCC  
CAGGCCTGTGCCCTCTCTGGTCTCGTGCTTGTGGCAGCTCTGTCTTCAGTTCTGGGATA  
TGTGCCCCTGTGGATGCTTCATTCCAGCCTCAGGGAAGCCTGGCACCCACTGCCCCAAGCT  
GAGCCAGAGGAAGGCTGAGTACTTGGTTCACAGAAGGAGATACTGGGTGGGAAAAAGATG  
GGGCAAAGCGGTATGATGCCTGGCAAAGGGCCTGCATGGCTATCCTCATTGCTACCTAAT  
GTGCTTGCAAAGCTCCATGTTTCCTAACAGATTGAGACTCCTGGCCAGGTGTGGTGGCC  
CACACCTGTAATTCTAGCACTTTGGGAGGCCAAGGTGGGCAGATCACTTGAGGTCAGGAG  
TTCAAGACCAGCCTGGCCAACATGGTGAAACTCCATCTCTACTAAAAAATAAATAACA  
AAAATTAGCTGGGTGCGCTAGTGCTGCCTGTAATCTCATCTACTCGGGAGGCTAAGACA  
GGAGACTCTCACTTCAACCCAGGAGGTGGAGGTTGCGGTGAGCCAAGATTGTGCCTCTGC  
ACTCTAGCGTGGGTGACAGAGTAAGCGAGACTCCATCTCAAAAATAAATAATAATAAT  
TCAGACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAACTCATGCCTGTAATCC  
CAACATTTTGGGAGGCCAACGCAGGAGGATTGCTTGAGGTCTGGAGGTTTGAGACCAGCC  
TGGGCAACATAGAAAGACCCCATCTCTAAATAAATGTTTTAAAAAT

132/392

## FIGURE 126

MELPFVTHLFLPLVFLTGLCSPFNLDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGA  
PWDGPSGDRRGDVYRCPVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETGDGCG  
FMVS

**Important features:**

**Signal peptide:**

amino acids 1-22

**Cell attachment sequence:**

amino acids 70-73

**N-glycosylation site:**

amino acids 98-101

**Integrins alpha chain proteins:**

amino acids 67-81

133/392

## FIGURE 127

GAGAGGACGAGGTGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCCGGCTCCCGGAGCCAG  
CCCTTTCCTAACCCAACCAACCTAGCCCAGTCCCAGCCGCCAGCGCCTGTCCCTGTCAC  
GGACCCAGCGTTACCATGTCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTC  
CCTTCTGCTCCTGGTAACTTGGGTTTTTACTCCTGTAACAACCTGAAATAACAAGTCTTGC  
TACAGAGAATATAGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTATGC  
TGACTGGTGTGCTTTCAGTCAGATGTTGCATCCAATTTTGGAGGAAGCTTCCGATGTCAT  
TAAGGAAGAATTTCCAAATGAAAATCAAGTAGTGTTGCCAGAGTTGATTGTGATCAGCA  
CTCTGACATAGCCCAGAGATACAGGATAAGCAAATACCCAACCCCTCAAAATGTTTCGTAA  
TGGGATGATGATGAAGAGAGAATACAGGGGTGAGCGATCAGTGAAAGCATTGGCAGATTA  
CATCAGGCAACAAAAAGTGACCCCATTTCAAGAAATTCGGGACTTAGCAGAAATCACCAC  
TCTTGATCGCAGCAAAAGAAATATCATTTGGATATTTTGAGCAAAAGGACTCGGACAACTA  
TAGAGTTTTTTGAACGAGTAGCGAATATTTTGCATGATGACTGTGCCTTTCTTCTGCATT  
TGGGGATGTTTTCAAAACCGGAAAGATATAGTGGCGACAACATAATCTACAAACCACCAGG  
GCATTCTGCTCCGGATATGGTGTACTTGGGAGCTATGACAAATTTTGATGTGACTTACAA  
TTGGATTCAAGATAAATGTGTTCTTGTCCGAGAAATAACATTTGAAAATGGAGAGGA  
ATTGACAGAAGAAGGACTGCCTTTTCTCATACTCTTTCACATGAAAGAAGATACAGAAAG  
TTTAGAAATATTCCAGAAATGAAGTAGCTCGGCAATTAATAAGTGAAAAAGGTACAATAAA  
CTTTTACATGCCGATTGTGACAAATTTAGACATCCTCTTCTGCACATACAGAAACTCC  
AGCAGATTGTCCTGTAATCGCTATTGACAGCTTTAGGCATATGTATGTGTTTGGAGACTT  
CAAAGATGTATTAATTCCTGGAAACTCAAGCAATTCGTATTTGACTTACATTCTGGAAA  
ACTGCACAGAGAATTCCATCATGGACCTGACCCAACCTGATACAGCCCCAGGAGAGCAAGC  
CCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAACTAGCACCCAGTGAATA  
TAGGTATACTCTATTGAGGGATCGAGATGAGCTTTTAAAAACTTGAAAAACAGTTTGTAAG  
CCTTTCAACAGCAGCATCAACCTACGTGGTGGAAATAGTAAACCTATATTTTCATAATTC  
TATGTGTATTTTTATTTTGAATAAACAGAAAGAAATTTAAAAA  
AAAAAAAAAAAAAAAAAAAAA

134/392

**FIGURE 128**

MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRF  
SQMLHPIFEEASDVKEEFPNENQVVFARVDCDQHS DIAQRYRISKYPTLKLFRNGMMMK  
REYRGQRSVKALADYIRQOKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKDSDNRYRVFER  
VANILHDDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDK  
CVPLVREITFENGEEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHAD  
CDKFRHPLLHIQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREF  
HHGPDPTDTAPGEQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL

**Important features:**

**Signal peptide:**

amino acids 1-29

**Endoplasmic reticulum targeting sequence:**

amino acids 403-406

**Tyrosine kinase phosphorylation site:**

amino acids 203-211

**Thioredoxin family proteins:**

amino acids 50-66

135/392

**FIGURE 129**

GAGCAGGACGGAGCCATGGACCCCGCCAGGAAAGCAGGTGCCCAGGCCATGATCTGGACT  
GCAGGCTGGCTGCTGCTGCTGCTGCTTTCGCGGAGGAGCGCAGGCCCTGGAGTGCTACAGC  
TGCGTGACAGAAAGCAGATGACGGATGCTCCCCGAACAAGATGAAGACAGTGAAGTGCGCG  
CCGGGCGTGGACGTCTGCACCGAGGCCGTGGGGGCGGTGGAGACCATCCACGGACAATTCTC  
GCTGGCAGTGCGGGGTTGCGGTTTCGGGACTCCCCGGCAAGAATGACCGCGGCCTGGAT  
CTTCACGGGCTTCTGGCGTTTCATCCAGCTGCAGCAATGCGCTCAGGATCGCTGCAACGCC  
AAGCTCAACCTCACCTCGCGGGCGCTCGACCCGGCAGGTAATGAGAGTGCATACCCGCCC  
AACGGCGTGGAGTGCTACAGCTGTGTGGGCCTGAGCCGGGAGGCGTGCCAGGGTACATCG  
CCGCCGGTTCGTGAGCTGCTACAACGCCAGCGATCATGTCTACAAGGGCTGCTTCGACGGC  
AACGTACCTTGACGGCAGCTAATGTGACTGTGTCTTGCCTGTCCGGGGCTGTGTCCAG  
GATGAATTCTGCACTCGGGATGGAGTAACAGGCCCAGGGTTCACGCTCAGTGGCTCCTGT  
TGCCAGGGGTCCCGCTGTAACCTGACCTCCGCAACAAGACCTACTTCTCCCCTCGAATC  
CCACCCCTTGTCCGGCTGCCCCCTCCAGAGCCACGACTGTGGCCTCAACCACATCTGTC  
ACCACTTCTACCTCGGCCCCAGTGAGACCCACATCCACCACCAACCCATGCCAGCGCCA  
ACCAGTCAGACTCCGAGACAGGGAGTAGAACACGAGGCCTCCCGGGATGAGGAGCCAGG  
TTGACTGGAGGCGCCGCTGGCCACCAGGACCGCAGCAATTGAGGGCAGTATCCTGCAAAA  
GGGGGGCCCCAGCAGCCCCATAATAAAGGCTGTGTGGCTCCACAGCTGGATTGGCAGCC  
CTTCTGTTGGCCGTGGCTGCTGGTGTCTTACTGTGAGCTTCTCCACCTGGAAATTTCCCT  
CTCACCTACTTCTCTGGCCCTGGGTACCCCTCTTCTCATCACTTCCTGTTCCCACTG  
GACTGGGCTGGCCAGCCCCTGTTTTTCCAACATTTCCCAGTATCCCAGCTTCTGCTGC  
GCTGGTTTTGCGGCTTTGGGAAATAAAATACCGTTGTATATATTCTGCCAGGGGTGTTCTA  
GCTTTTTGAGGACAGCTCCTGTATCCTTCTCATCCTTGTCTCTCCGCTTGTCTCTTGTG  
ATGTTAGGACAGAGTGAGAGAAGTCAGCTGTACGGGGAAGGTGAGAGAGAGGATGCTAA  
GCTTCCTACTCACTTTCTCCTAGCCAGCCTGGACTTTGGAGCGTGGGGTGGGTGGGACAA  
TGGCTCCCCACTCTAAGCACTGCCTCCCCTACTCCCCGCATCTTTGGGGAATCGGTTCCC  
CATATGTCTTCTTACTAGACTGTGAGCTCCTCGAGGGGGGGCCCGGTACCCAATTCGCC  
CTATAGTGAGTCGTA

136/392

**FIGURE 130**

MDPARKAGAQAMIWTAGWLLLLLLRGGAQALECYSCVQKADDGCS PNKMKT VKCAPGVDV  
CTEAVGAVETIHGQFSLAVRGCGSGLPGKNDRGLDLHGLLAFIQLQQCAQDRCNAKLNLT  
SRALDPAGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGCDFDGNVTLT  
AANVTVSLPVRGCVQDEFCTRDGVTGPGFTLSGCCQGSRCNSDLRNKTYFSPRIPLVR  
LPPPEPTTVASTTSVTTSTSAFVRPTSTTKPMPAPTSQTPRQGEHEASRDEEPRLTGGA  
AGHQDRSNSGQYPAKGGPQQPHNKGCVAPTAGLAALLLAVAAGVLL



137/392

**FIGURE 131**

AAACTTGACGCCATGAAGATCCCGGTCTTCCTGCCGTGGTGCTCCTCTCCCTCCTGGTG  
CTCCACTCTGCCCAGGGAGCCACCCTGGGTGGTCCTGAGGAAGAAAGCACCATTGAGAAT  
TATGCGTCACGACCCGAGGCCTTTAACACCCCGTTCCTGAACATCGACAAATTGCGATCT  
GCGTTTAAGGCTGATGAGTTCCTGAACTGGCACGCCCTCTTTGAGTCTATCAAAAGGAAA  
CTTCCTTTCCTCAACTGGGATGCCTTTCCTAAGCTGAAAGGACTGAGGAGCGCAACTCCT  
GATGCCCAGTGACCATGACCTCCACTGGAAGAGGGGGCTAGCGTGAGCGCTGATTCTCAA  
CCTACCATAACTCTTTCCTGCCTCAGGAAGTCCAATAAAACATTTTCCATCCAAA

138/392

## FIGURE 132

MKIPVLPVVLSSLLVLHSAQGATLGGPEEESTIENYASRPEAFNTPFLNIDKLRSAFKA  
DEFLNWHALFESIKRKLPFLNWDAFPKLKGLRSATPDAQ

139/392

**FIGURE 133**

CAGTTCTGAAATCAATGGAGTTAATTTAGGGAATACAAACCAGCCATGGGGGTGGAGATT  
GCCTTTGCCTCAGTGATTCTCACCTGCCTCTCCCTTCTGGCAGCAGGAGTCTCCAGGTT  
GTTCTTCTCCAGCCAGTTCCAACCTCAGGAGACAGGTCCCAAGGCCATGGGAGATCTCTCC  
TGTGGCTTTGCCGGCCACTCATGAGAGTGTTTTTGTGTAAAGTATTTTTTTAGAATACTGT  
TGA~~CTT~~CTTCATGATTTAATAACCATCCTTTGCGAAGTTTTATGAGGCTTTAGGGGAATG  
TCAACCCTCAAATTTTTGTTATACTAGATGGCTTCCATTTACCCACCACTATTTTAAGGT  
CCCTTTATTTTTAGGTTCAAGGTTCA~~TTT~~GACTTGAGAAAGTGCCCTTCTGCAGCTTCAT  
TGATTTTGTTTATCTTCACTATTAATTGTAACGATTAAAAAAGAATAAGAGCACGCAGAC  
CTCTAGGAGAATATTTTATCCCTGGGTGCCCCTGACACATTTATGTAGTGATCCACAAA  
TGTGATTGTTAATTTAAATGTTATTCTAATATTAGTACATTCAGTTGTGATGTAATATGA  
ATAACCAGAATCTATTTCTTAAAAGTTTTGAGTATATTTTTTCAACTAGATATTTGTATAG  
AAAGACTGAATAGTGATG

140/392

## FIGURE 134

MGVEIAFASVILTCLSLAAGVSQVVLLQPVPQTQETGPKAMGDLSCGFAGHS

141/392

**FIGURE 135**

GGGGAATCTGCAGTAGGTCTGCCGGCGATGGAGTGGTGGGCTAGCTCGCCGCTTCGGCTC  
TGGCTGCTGTTGTTCCCTCCTGCCCTCAGCGCAGGGCCGCCAGAAGGAGTCAGGTTCAAAA  
TGGAAAGTATTTATTGACCAAATTAACAGGTCTTTGGAGAATTACGAACCATGTTCAAGT  
CAAACTGCAGCTGCTACCATGGTGTCATAGAAGAGGATCTAACTCCTTTCCGAGGAGGC  
ATCTCCAGGAAGATGATGGCAGAGGTAGTCAGACGGAAGCTAGGGACCCACTATCAGATC  
ACTAAGAACAGACTGTACCGGGAAAATGACTGCATGTTCCCCTCAAGGTGTAGTGGTGTT  
GAGCACTTTATTTTGGAAAGTGATCGGGCGTCTCCCTGACATGGAGATGGTGATCAATGTA  
CGAGATTATCCTCAGGTTCTAAATGGATGGAGCCTGCCATCCCAGTCTTCTCCTTCAGT  
AAGACATCAGAGTACCATGATATCATGTATCCTGCTTGGACATTTTGGGAAGGGGGACCT  
GCTGTTTTGGCCAATTTATCCTACAGGTCTTGGACGGTGGGACCTCTTCAGAGAAGATCTG  
GTAAGGTCAGCAGCACAGTGGCCATGGAAAAAGAAAACTCTACAGCATATTTCCGAGGA  
TCAAGGACAAGTCCAGAACGAGATCCTCTCATTCTTGTCTCGGAAAAACCCAAACTT  
GTTGATGCAGAATACACCAAAAACCAGGCCTGGAAATCTATGAAAGATACCTTAGGAAAG  
CCAGCTGCTAAGGATGTCCATCTTGTGGATCACTGCAAATACAAGTATCTGTTTAATTTT  
CGAGGCGTAGCTGCAAGTTTCCGGTTTAAACACCTCTTCCTGTGTGGCTCACTTGTTTTT  
CATGTTGGTGATGAGTGGCTAGAATTCTTCTATCCACAGCTGAAGCCATGGGTTCACTAT  
ATCCCAGTCAAAACAGATCTCTCCAATGTCCAAGAGCTGTTACAATTTGTAAAAGCAAAT  
GATGATGTAGCTCAAGAGATTGCTGAAAGGGGAAGCCAGTTTATTAGGAACCATTTGCAG  
ATGGATGACATCACCTGTTACTGGGAGAACCTCTTGAGTGAATACTCTAAATTCCTGTCT  
TATAATGTAACGAGAAGGAAAGGTTATGATCAAATTATCCCAAATGTTGAAAAC TGAA  
CTATAGTAGTCATCATAGGACCATAGTCCTCTTTGTGGCAACAGATCTCAGATATCCTAC  
GGTGAGAAGCTTACCATAAGCTTGGCTCCTATACCTTGAATATCTGCTATCAAGCCAAAT  
ACCTGGTTTTTCCTTATCATGCTGCACCCAGAGCAACTCTTGAGAAAGATTTAAAATGTGT  
CTAATACACTGATATGAAGCAGTTCAACTTTTTTGGATGAATAAGGACCAGAAATCGTGAG  
ATGTGGATTTTGAACCCAACCTCTACCTTTTCAATTTCTTAAGACCAATCACAGCTTGTGCC  
TCAGATCATCCACCTGTGTGAGTCCATCACTGTGAAATTGACTGTGTCCATGTGATGATG  
CCCTTTGTCCCATTTATTTGGAGCAGAAAATTCGTCAATTTGGAAGTAGTACAACCTCATTGC  
TGGAATTGTGAAATTATTCAAGGCGTGATCTCTGTCACTTTATTTTAATGTAGGAAACCC  
TATGGGGTTTATGAAAAATACTTGGGGATCATTCTCTGAATGGTCTAAGGAAGCGGTAGC  
CATGCCATGCAATGATGTAGGAGTTCTCTTTTGTAACCATAAACTCTGTTACTCAGGA  
GGTTTCTATAATGCCACATAGAAAGAGGCCAATTGCATGAGTAATTATTGCAATTGGATT  
TCAGGTTCCCTTTTTTGTGCCTTCATGCCCTACTTCTTAATGCCTCTCTAAAGCCAAA

142/392

**FIGURE 136**

MEWWASSPLRLWLLLFLPSAQGRQKESGSKWKV FIDQINRSLENYEPCSSQNCSCYHGV  
IEEDLTPFRGGISRKMAEVVRRKLGTHYQITKNRLYREND CMFPSRCSGVEHFILEVIG  
RLPDMEMVINVRDYPQVPKWMEPAIPVFSFSKTSEYHDIMYPAWTFWEGGPAVWPIYPTG  
LGRWDLFREDLVRSAQWPWKKKNSTAYFRGSRTSPERDPLILLSRKNPKLVDAEYTKNQ  
AWKSMKDTLGKPAAKDVHLVDHCKYKYLNFNRGVAASFRFKHLFLCGSLVFHVGDEWLEF  
FYPQLKPWVHYIPVKTDLSNVQELLQFVKANDDVAQEIAERGSQFIRNHLQMDDITCYWE  
NLLSEYSKFLSYNVTRRKG YDQIIPKMLKTEL

143/392

**FIGURE 137**

ATTCTCCTAGAGCATCTTTGGAAGCATGAGGCCACGATGCTGCATCTTGGCTCTTGTCTG  
CTGGATAACAGTCTTCCTCCTCCAGTGTTCAAAGGAACTACAGACGCTCCTGTTGGCTC  
AGGACTGTGGCTGTGCCAGCCGACACCCAGGTGTGGGAACAAGATCTACAACCCCTTCAGA  
GCAGTGCTGTTATGATGATGCCATCTTATCCTTAAAGGAGACCCGCCGCTGTGGCTCCAC  
CTGCACCTTCTGGCCCTGCTTTGAGCTCTGCTGTCCCGAGTCTTTTGGCCCCCAGCAGAA  
GTTTCTTGTGAAGTTGAGGGTTCTGGGTATGAAGTCTCAGTGTCACTTATCTCCCATCTC  
CCGGAGCTGTACCAGGAACAGGAGGCACGTCCTGTACCCATAAAAACCCAGGCTCCACT  
GGCAGACGGCAGACAAGGGGAGAAGAGACGAAGCAGCTGGACATCGGAGACTACAGTTGA  
ACTTCGGAGAGAAGCAACTTGACTTCAGAGGGATGGCTCAATGACATAGCTTTGGAGAGG  
AGCCCAGCTGGGGATGGCCAGACTTCAGGGGAAGAATGCCTTCCTGCTTCATCCCCTTTC  
CAGCTCCCCTTCCCGCTGAGAGCCACTTTCATCGGCAATAAAATCCCCACATTTACCATCT

144/392

## FIGURE 138

MRPRCCILALVCWITVFLQCSKGTDDAPVGSGLWLCQPTPRCGNKIYNPSEQCCYDDAI  
LSLKETRRCGSTCTFWPCFELCCPESFGPQQKFLVKLRVLGMKSQCHLSPISRSCTRNR  
HVLYP

**Important features:**

**Signal sequence:**

amino acids 1-21

**N-myristoylation sites:**

amino acids 33-39, 70-76



145/392

**FIGURE 139**

CCTCTGTCCACTGCTTTTCGTGAAGACAAGATGAAGTTCACAATTGTCTTTGCTGGACTTC  
TTGGAGTCTTCTAGCTCCTGCCCTAGCTAACTATAATATCAACGTCAATGATGACAACA  
ACAATGCTGGAAGTGGGCAGCAGTCAGTGAGTGTCAACAATGAACACAATGTGGCCAATG  
TTGACAATAACAACGGATGGGACTCCTGGAATTCATCTGGGATTATGGAAATGGCTTTG  
CTGCAACCAGACTCTTTCAAAAGAAGACATGCATTGTGCACAAAATGAACAAGGAAGTCA  
TGCCCTCCATTCAATCCCTTGATGCACTGGTCAAGGAAAAGAAGCTTCAGGGTAAGGGAC  
CAGGAGGACCACCTCCCAAGGGCCTGATGTACTCAGTCAACCCAAACAAAGTCGATGACC  
TGAGCAAGTTCGGAAAAAACATTGCAAACATGTGTCTGTGGGATTCCAACATACATGGCTG  
AGGAGATGCAAGAGGCAAGCCTGTTTTTTTACTCAGGAACGTGCTACACGACCAGTGTAC  
TATGGATTGTGGACATTTCTTCTGTGGAGACACGGTGGAGAACTAAACAATTTTTTAAA  
GCCACTATGGATTTAGTCATCTGAATATGCTGTGCAGAAAAAATATGGGCTCCAGTGGTT  
TTTACCATGTCATTCTGAAATTTTCTCTACTAGTTATGTTTGATTCTTTAAGTTTCAA  
TAAAATCATTTAGCATTGAAAAAA

146/392

**FIGURE 140**

MKFTIVFAGLLGVFLAPALANYNINVNDDNNNAGSGQQSVSVNNEHNVANVDNNNGWDSW  
NSIWYDYGNGFAATRLFQKKTCIVHKMNKEVMPSIQSLDALVKEKKLQGKGPGGPPPKGLM  
YSVNPKNKVDLDSKFGKNIANMCRGIPTYMAEEMQEASLFFYSGTCYTTSVLWIVDISFCG  
DTVEN

**Signal Peptide:**  
amino acids 1-20

**N-myristoylation Sites:**  
amino acids 67-72, 118-123, 163-168

**Flavodoxin protein homology:**  
amino acids 156-174

147/392

**FIGURE 141**

GGTCCTTAATGGCAGCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCC  
TGCTGTCCGGCTGGTCCCGGGCTGGGCGAGCCGACCCTCACTCTCTTTGCTATGACATCA  
CCGTCATCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATG  
AAAAGACTTTTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGG  
GGAAGAACTAAATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGG  
TGGACATACTTACAGAGCAACTGCGTGACATTAGCTGGAGAATTACACACCCAAGGAAC  
CCCTCACCTGTCAGGCAAGGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGAT  
CTTGGCAGTTCAGTTTCGATGGGCAGATCTTCCTCCTCTTTGACTCAGAGAAGAGAATGT  
GGACAACGGTTCATCCTGGAGCCAGAAAGATGAAAGAAAAGTGGGAGAATGACAAGGTTG  
TGGCCATGTCCTTCCATTACTTCTCAATGGGAGACTGTATAGGATGGCTTGAGGACTTCT  
TGATGGGCATGGACAGCACCTGGAGCCAAGTGCAGGAGCACCCTCGCCATGTCCTCAG  
GCACAACCCAACTCAGGGCCACAGCCACCACCCTCATCCTTTGCTGCCTCCTCATCATCC  
TCCCCTGCTTCATCCTCCCTGGCATCTGAGGAGAGTCCTTTAGAGTGACAGGTTAAAGCT  
GATACCAAAGGCTCCTGTGAGCACGGTCTTGATCAAACCTCGCCCTTCTGTCTGGCCAGC  
TGCCCACGACCTACGGTGTATGTCCAGTGGCCTCCAGCAGATCATGATGACATCATGGAC  
CCAATAGCTCATTCACTGCCTTGATTCTTTTGCCAACAATTTTACCAGCAGTTATACCT  
AACATATTATGCAATTTTCTCTTGGTGCTACCTGATGGAATTCCTGCACTTAAAGTTCTG  
GCTGACTAAACAAGATATATCATTTTCTTCTTCTCTTTTGTGTTGGAAAATCAAGTACT  
TCTTTGAATGATGATCTCTTTCTTGCAAATGATATTGTCAGTAAAATAATCACGTTAGAC  
TTCAGACCTCTGGGGATTCTTTCCGTGTCCTGAAAGAGAATTTTAAATTATTTAATAAG  
AAAAAATTTATATTAATGATTGTTTCCTTTAGTAATTTATTGTTCTGTACTGATATTTAA  
ATAAGAGTTCTATTTCCCAAAAAAAAAAAAAAAAAAAAA

148/392

**FIGURE 142**

MAAAAATKILLCLPLLLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKT  
FLHYDCGNKTVTPVSPLGKKLNVTTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLT  
LQARMSCEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAM  
SFHYFSMGDCIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLLILCCLLIILPC  
FILPGI

**Important features:****Signal peptide:**

amino acids 1-25

**Transmembrane domain:**

amino acids 224-246

**N-glycosylation site:**

amino acids 68-72, 82-86

**N-myristoylation site:**

amino acids 200-206, 210-216

**Amidation site:**

amino acids 77-81

149/392

**FIGURE 143**

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTTAGCACCAGTACTGG  
ATGTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTCAGTGTCCGATTCTGATTCCGGC  
AAGGATCCAAGCATGGGAATGCTGCCGTCGGGCAACTCCTGGCACACTGCTCCTCTTCTG  
GCTTTCCTGCTCCTGAGTTCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGG  
GATGCCTGGGGCCCATGGAGTGAATGCTCACGCACCTGCCGGGGAGGGGCCTCCTACTCT  
CTGAGGCGCTGCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGC  
AGTAATGTGGACTGCCCACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCAT  
AATGATGTCAAGCACCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGAC  
AACCCATGTTCACTCAAGTGCCAAGCCAAAGGAACAACCCTGGTTGTTGAAGTAGCACCT  
AAGGTCTTAGATGGTACGCGTTGCTATACAGAACTTTTGGATATGTGCATCAGTGGTTTA  
TGCCAAATTGTTGGCTGCGATCACCAGCTGGGAAGCACCGTCAAGGAAGATAAATCCCAGCT  
GTCTGCAACGGAGATGGGTCCACCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCT  
TCCGCAACCAAATCGGATGATACTGTGGTTGCACTTCCCTATGGAAGTAGACATATTCGC  
CTTGTCTTAAAGGTCTGATCACTTATATCTGGAAACCAAAACCCTCCAGGGGACTAAA  
GGTGAACACAGTCTCAGCTCCACAGGAACCTTCTTGTGGACAATTCTAGTGTGGACTTC  
CAGAAATTTCCAGACAAAGAGATACTGAGAATGGCTGGACCACTCACAGCAGATTTTCAAT  
GTCAAGATTTCGTAACTCGGGCTCCGCTGACAGTACAGTCCAGTTCATCTTCTATCAACCC  
ATCATCCACCGATGGAGGGAGACGGATTTCTTTCCTTGCTCAGCAACCTGTGGAGGAGGT  
TATCAGCTGACATCGGCTGAGTGCTACGATCTGAGGAGCAACCGTGTGGTTGCTGACCAA  
TACTGTCACTATTACCCAGAGAACATCAAACCCAAACCCAAGCTTCAGGAGTGCAACTTG  
GATCCTTGTTCCAGCCAGTGACGGATACAAGCAGATCATGCCTTATGACCTCTACCATCCC  
CTTCTCGGTGGGAGGCCACCCCATGGACCGCGTGCTCCTCCTCGTGTGGGGGGGGCATC  
CAGAGCCGGGCAGTTTCTGTGTGGAGGAGGACATCCAGGGGCATGTCACTTCAGTGGAA  
GAGTGGAAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCCTGCAACATTTTTGAC  
TGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGTGGCCAGGGCCTC  
AGATAACCGTGTGGTCTCTGCATCGACCATCGAGGAATGCACACAGGAGGCTGTAGCCCA  
AAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATAAACCCAAA  
GAGAACTTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTAGAAGAA  
GGAGCTGCTGTGTGTCAGAGGAGCCCTCGTTAAGTTGTAAAAGCACAGACTGTTCTATATTG  
AAACTGTTTTGTTTAAAGAAAGCAGTGTCTCACTGGTTGTAGCTTTCATGGGTTCTGAAC  
TAAGTGTAATCATCTCACCAAAGCTTTTTGGCTCTCAAATTAAAGATTGATTAGTTTCAA  
AAAAAAA

150/392

**FIGURE 144**

MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTC GGGASYS LRRC  
LSSKSCEGRNIRYRTC SNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCS  
LKCQAKGTTLVVELAPKVLDTGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNG  
DGSTCRLVRGQYKSQLSATKSDDTVVALPYGSRHRLVLKGPDLHLYLETKTLQGTKGENS  
LSSTGTFLVDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHR  
WRETDFFPCSATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCP  
ASDGYKQIMPYDLYHPLPRWEATPWTACSSSCGGGIQSRVSCVEEDIQGHVTSVEEWKC  
MYTPKMPIAQPCNIFDCPKWLAQEWS PCTVT CGQGLRYRVVLCIDHRGMHTGGCSPKTKP  
HIKEECIVPTPCYKPKKEKLPVEAKLPWFKQAQELEEGA AVSEEPS

**Important features:**

**Signal peptide:**

amino acids 1-25

**N-glycosylation site:**

amino acids 251-254

**Thrombospondin 1:**

amino acids 385-399

**von Willebrand factor type C domain proteins:**

amino acids 385-399, 445-459 and 42-56

151/392

**FIGURE 145**

GGAGGAGGGAGGGCGGGCAGGCGCCAGCCCAGAGCAGCCCCGGGCACCAGCACGGACTCT  
CTCTTCCAGCCCAGGTGCCCCCACTCTCGCTCCATTTCGGCGGGAGCACCCAGTCCTGTA  
CGCCAAGGAACTGGTCTCTGGGGGCACCATGGTTTCGGCGGCAGCCCCCAGCCTCCTCATC  
CTTCTGTTGCTGCTGCTGGGGTCTGTGCCTGCTACCGACGCCCCGCTCTGTGCCCCCTGAAG  
GCCACGTTCTTGAGGATGTGGCGGGTAGTGGGGAGGCCGAGGGCTCGTCGGCCTCCTCC  
CCGAGCCTCCCGCCACCCTGGACCCCCGGCCCTCAGCCCCACATCGATGGGGCCCCAGCCC  
ACAACCCTGGGGGGCCCATCACCCCCACCAACTTCCTGGATGGGATAGTGGACTTCTTC  
CGCCAGTACGTGATGCTGATTGCTGTGGTGGGCTCCCTGGCCTTTCTGCTGATGTTTCATC  
GTCTGTGCCGCGGTATCACCCGGCAGAAGCAGAAGGCCTCGGCCTATTACCCATCGTCC  
TTCCCCAAGAAGAAGTACGTGGACCAGAGTGACCGGGCCGGGGGCCCCCGGGCCTTCAGT  
GAGGTCCCCGACAGAGCCCCCGACAGCAGGCCCCGAGGAAGCCCTGGATTCTCCCGGCAG  
CTCAGGCCGACATCTTGGCCGCCACCCAGAACCTCAAGTCCCCCACCAGGGCTGCACTG  
GGCGGTGGGGACCGAGCCAGGATGGTGGAGGGCAGGGGCGCAGAGGAAGAGGAGAAGGGC  
AGCCAGGAGGGGGACAGGAAGTCCAGGGACATGGGGTCCAGTGGAGACACCAGAGGCG  
CAGGAGGAGCCGTGCTCAGGGGTCTTTAGGGGGCTGTGGTGGCCGGTGAGGGCCAAGGG  
GAGCTGGAAGGGTCTCTCTTGTAGCCCAGGAAGCCCAGGGACCAGTGGGTCCCCCGAA  
AGCCCCCTGTGCTTGCAGCAGTGTCCACCCCAGTGTCTAAACAGTCTCTCCCGGGCTGCCAGC  
CCTGACTGTGCGGGCCCCCAAGTGGTCACCTCCCCGTGTATGAAAAGGCCTTCAGCCCTGA  
CTGCTTCTGACACTCCCTCCTTGGCCTCCCTGTGGTGCCAATCCCAGCATGTGCTGATT  
CTACAGCAGGCAGAAATGCTGGTCCCCGGTGCCCCGGAGGAATCTTACCAAGTGCCATCA  
TCCTTCACCTCAGCAGCCCCAAAGGGCTACATCCTACAGCACAGTCCCCCTGACAAAGTG  
AGGGAGGGCACGTGTCCCTGTGACAGCCAGGATAAAACATCCCCAAAGTGCTGGGATTA  
CAGGCGTGAGCCACCGTGCCCCGGCCCCAACTACTTTTTTAAACAGCTACAGGGTAAATC  
CTGCAGCACCCACTCTGGAAAATACTGCTCTTAATTTTCTGAAGGTGGCCCCCTGTTTC  
TAGTTGGTCCAGGATTAGGGATGTGGGGTATAGGGCATTTAAATCCTCTCAAGCGCTCTC  
CAAGCACCCCCGGCCTGGGGGTGAGTTTCTCATCCCGCTACTGCTGCTGGGATCAGGTTG  
AATGAATGGAATCTTCTGTCTGGCCTCCAAAGCAGCCTAGAAGCTGAGGGGCTGTGTT  
TGAGGGGACCTCCACCCTGGGGAAAGTCCGAGGGGCTGGGGAAAGGTTTCTGACGCCCAGC  
CTGGAGCAGGGGGGCCCTGGCCACCCCCTGTTGCTCACACATTGTCTGGCAGCCTGTGTC  
CACAAATATTCGTCAGTCTCGACAGGGAGCCTGGGCTCCGTCTGCTTTAGGGAGGCTCT  
GGCAGGAGGTCTCTCCCCATCCCTCCATCTGGGGCTCCCCAACCTCTGCACAGCTCT  
CCAGGTGCTGAGATATAATGCACCAGCACATAAACCTTTATTCCGGCCTGAAAAAAAAA  
AAGA

152/392

**FIGURE 146**

MVSAAAPSLILLILLLLGSPATDARSVPLKATFLEDVAGSGEAGSSASSPSLPPPWTP  
ALSPTSMGPQPTTLGGPSPPTNFDGIVDFFRQYVMLIAVVGSLAFLLMFIVCAAVITRQ  
KQKASAYYPSSFPPKKKYVDQSDRAGGPRAFSEVPDRAPDSRPEEALDSSRQLQADILAAAT  
QNLKSPTRAALGGGDGARMVEGRGAEEEEKGSQEGDQEVQGHGVPVETPEAQEEPCSGVL  
EGAVVAGEGQGELEGSLLLAQEAQGPVGPPEPCACSSVHPSV

**Signal peptide:**

amino acids 1-25

**Transmembrane domain:**

amino acids 94-118

**N-myristoylation site:**amino acids 18-24, 40-46, 46-52, 145-151, 192-198, 193-199,  
211-217, 238-244, 242-248



153/392

**FIGURE 147**

GAAAGACGTGGTCCTGACAGACAGACAATCCTATTCCCTACCAAAATGAAGATGCTGCTG  
CTGCTGTGTTTGGGACTGACCCTAGTCTGTGTCCATGCAGAAGAAGCTAGTTCTACGGGA  
AGGAACTTTAATGTAGAAAAGATTAATGGGGAATGGCATACTATTATCCTGGCCTCTGAC  
AAAAGAGAAAAGATAGAAGAACATGGCACTTTTAGACTTTTTCTGGAGCAAATCCATGTC  
TTGGAGAATTCCTTAGTTCTTAAAGTCCATACTGTAAGAGATGAAGAGTGCTCCGAATTA  
TCTATGGTTGCTGACAAAACAGAAAAGGCTGGTGAATATTCTGTGACGTATGATGGATT  
AATACATTTACTATACCTAAGACAGACTATGATAACTTTCTTATGGCTCACCTCATTAAC  
GAAAAGGATGGGGAAACCTTCCAGCTGATGGGGCTCTATGGCCGAGAACCAGATTTGAGT  
TCAGACATCAAGGAAAGGTTTGCACAACTATGTGAGGAGCATGGAATCCTTAGAGAAAAT  
ATCATTGACCTATCCAATGCCAATCGCTGCCTCCAGGCCCGAGAATGAAGAATGGCCTGA  
GCCTCCAGTGTTGAGTGGACACTTCTCACCAGGACTCCACCATCATCCCTTCTTATCCAT  
ACAGCATCCCCAGTATAAATTCTGTGATCTGCATTCCATCCTGTCTCACTGAGAAGTCCA  
ATTCCAGTCTATCAACATGTTACCTAGGATACCTCATCAAGAATCAAAGACTTCTTTAAA  
TTTCTCTTTGATACACCCTTGACAATTTTTTCATGAAATTATTCCTCTTCCTGTTCAATAA  
ATGATTACCCTTGCACTTAA

154/392

## FIGURE 148

MKMLLLLCLGLTLVCVHAEASSTGRNFNVEKINGEWHTIILASDKREKIEEHGNFRLFL  
EQIHVLENSLVLKVHTVRDEECSELSMVADKTEKAGEYSVTYDGFNTFTIPKTDYDNFLM  
AHLINEKDGETFQLMGLYGREPDLSSEDIKERFAQLCEEHGILRENIIDLSNANRCLQARE

155/392

**FIGURE 149**

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGG  
GAGGGAGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAG  
GGCGGCAAGGAGGAGACCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCA  
GAGATGAAGTTCCAGGGGCCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGG  
GAGGCTGGCCCCCTGCAGAGCGGAGAGGAAAAGCACTGGGACAAATATTGGGGAGGCCCTT  
GGACATGGCCTGGGAGACGCCCTGAGCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCC  
GGAGGGGCAGCTGGCTCTAAAGTCAGTGAGGCCCTTGGCCAAGGGACCAGAGAAGCAGTT  
GGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGCAGCAGATGCTTTGGGCAACAGGGTC  
GGGGAAGCAGCCCATGCTCTGGGAAACACTGGGCACGAGATTGGCAGACAGGCAGAAGAT  
GTCATTTCGACACGGAGCAGATGCTGTCCGCGGCTCCTGGCAGGGGGTGCCTGGCCACAGT  
GGTGCTTGGGAAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGGCCTTGGAGGC  
CAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATACCCCGGAAAC  
TCAGCAGGCAGCTTTGGAATGAATCCTCAGGGAGCTCCCTGGGGTCAAGGAGGCAATGGA  
GGGCCACCAAACCTTTGGGACCAACACTCAGGGAGCTGTGGCCCAGCCTGGCTATGTTCA  
GTGAGAGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCACCATCTGGCTCAGGTGGA  
GGCTCCAGCAACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGC  
AATGGTGACAACAACAATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGC  
AGCAGCAGTGGCGGCAGCAGTGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGC  
AGCAGAGGTGACAGCGGCAGTGAGTCCTCCTGGGGATCCAGCACCGGCTCCTCCTCCGGC  
AACCACGGTGGGAGCGGCGGAGGAAATGGACATAAACC CGGTGTGAAAAGCCAGGGAAT  
GAAGCCCGCGGAGCGGGGAATCTGGGATTTCAGGGCTTCAGAGGACAGGGAGTTTCCAGC  
AACATGAGGGAAATAAGCAAGAGGGCAATCGCCTCCTTGGAGGCTCTCGAGACGATTTAT

156/392

**FIGURE 150**

MKFQGPLACLLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEG VGKAIGKEAG  
 GAAGSKVSEALGQGTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDV  
 IRHGADAVRGSWQGVPGHSGAWETSGGHGIFGSQGGGLGGQGQGNPGLGTPVWHGYPGNS  
 AGSFGMNPQGAPWGQGGNGGPPNFGTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGG  
 SSNSGGGSGSGSGSSGSGSNGDNNNGSSSSGGSSSGSSSGSSSGSSSGSGSSSGNSGGS  
 RGDSGSESSWGSSTGSSSGNHGSGGGNGHKPGCEKPGNEARGSGESGIQGFRQGVSSN  
 MREISKEGNRLLGGSGDNYRGQSSWGSGGGDAVGGVNTVNSETPGMFNFDTFWKNFKS  
 KLGFINWDAINKDORSSRIP

Signal peptide:  
amino acids 1-21

N-glycosylation site:  
amino acids 265-269

Glycosaminoglycan attachment site:  
amino acids 235-239, 237-241, 244-248, 255-259, 324-328,  
388-392

**Casein kinase II phosphorylation site:**  
amino acids 26-30, 109-113, 259-263, 300-304, 304-308

**N-myristoylation site:**  
amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70,  
74-80, 90-96, 96-102, 130-136, 140-146, 149-155, 152-158,  
155-161, 159-165, 163-169, 178-184, 190-196, 194-200,  
199-205, 218-224, 236-242, 238-244, 239-245, 240-246,  
245-251, 246-252, 249-252, 253-259, 256-262, 266-272,  
270-276, 271-277, 275-281, 279-285, 283-289, 284-290,  
287-293, 288-294, 291-297, 292-298, 295-301, 298-304,  
305-311, 311-317, 315-321, 319-325, 322-328, 323-329,  
325-331, 343-349, 354-360, 356-362, 374-380, 381-387,  
383-389, 387-393, 389-395, 395-401

Cell attachment sequence:  
amino acids 301-304

157/392

**FIGURE 151**

CGGCCACAGCTGGCATGCTCTGCCTGATCGCCATCCTGCTGTATGTCTCGTCCAGTACC  
TCGTGAACCCCGGGGTGCTCCGCACGGACCCAGATGTCAAGAATATGAACACGTGGCTG  
CTGTTCTCTCCCTGTTCCCGGTGCAGGTGCAGACCCTGATAGTCGTGATCATCGGGATG  
CTCGTGCTCCTGCTGGACTTTCTTGGCTTGGTGCACCTGGGCCAGCTGCTCATCTTCCAC  
ATCTACCTGAGTATGTCCCCACCCCTAAGCCCCGATCCCCCAAGGCTGGGTGGTGCAGA  
GCTGCTCATCTTACACCTCTACTTGAGTATGTCCCTAACCCTGAGCCCCCACGCCTGGG  
GCCAGAGTCTTTGTCCCCGTGTGCGCATGTGTTCAAGGTGAGCCTCTCCAGAAGTGAG  
ATCATGGACAAAAAGGGCAAATCACAGGAAGAAATTAAATCCATGAGGACCCAGCAGGCC  
CAGCAAGAAGCTGAACTCACGCCGAGACCTGCAGGAGTGGTGCCAGGTGCTTGAAGTAAC  
AAGTTTAAATGTTTACAGAGACAATGGAATGGAATCTATTAGGCAAGAACAGGACATTATG  
AAATAAGGACAGGTGGACTTCCAAAAACACAAGTAGAAATTCTAACAATGAAATATATTA  
CAGGCAGGTCACCCACTAACCAACAACCTGAAGCGAGAGCTGTGGTCTTGCTTGGTCTCA  
CAGTGGGCACAGCGGTAGGCGGTGAGTCATGTTGCTGAACGACGGAGGGTAACTCCCCA  
GCCCCAAGAAAACCTGTGTTGGAAGTAACAACAACCTCCCTGCTCCTGGCACCAGCCGTT  
TTGGTCATGGTGGGCCAGCTGCAAAGCGTCTTCCATTCTCTGGGCAGTGGTGGCCCCGAG  
GCTGTGGCCTCTCAGGGGGTTTCTGTGGACACGGGCAGCAGAGTGTGTCCAGGCCAGCCC  
CCAAGAATGCCCTGCTCCTGACAGCTTGGCCAACCCCTGGTCAGGGCAGAGGGAGTTGGG  
TGGGTGAGGCTCTGGGCTCACCTCCATCTCCAGAGCATCCCCTGCCTGCAGTTGTGGCAA  
GAACGCCAGCTCAGAATGAACACACCCCAAGAGCCTCCTTGTTCATAACCACAGGT  
TACCCTACAAACCACTGTCCCCACACAACCTGGGGATGTTTTAAACACACACCTCTAA  
CGCATATCTTACAGTCACTGTTGTCTTGCTGAGGGTTGAATTTTTTTTAAATGAAAGTGC  
AATGAAAATCACTGGATTAAATCCTACGGACACAGAGCTGAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAA

158/392

## FIGURE 152

MNTWLLFLPLFPVQVQTLIVVIIGMLVLLLDLGLVHLGQLLIFHIYLSMSPTLSPRSPQ  
GWVVRAAHLTPLEYPNPEPPTPGARVFVPRVRMCSGSASPRSEIMDKKGKSQEEIKSM  
RTQQAQQEALTPRPAGVVPGA

159/392

**FIGURE 153**

AACTGGAAGGAAAGAAAGAAAGGTCAGCTTTGGCCCAGATGTGGTTACCCCTTGGTCTCC  
TGTCTTTATGTCTTTCTCCTCTTCCTATTCTGTTCATCTCCCTCACTTAAGTCTCAGGCCT  
GTCAGCAGCTCCTGTGGACATTGCCATCCCCCTCTGGTAGCCTTCAGAGCAAACAGGACAA  
CCTATGTTATGGATGTTTCCACCAACCAGGGTAGTGGCATGGAGCACCGTAACCATCTGT  
GCTTCTGTGATCTCTATGACAGAGCCACTTCTCCACCTCTGAAATGTTCCCTGCTCTGAA  
ATCTGGCATGAGATGGCACAGGTGACCACGCAGAAGCCACCAGAATCTTGCCTGCCCTAT  
TCCTCCTCCCAAGTCTGTTCTCTTATTGTCAACCTCAGCACAAACAGGCTGGCGCCAATGG  
CATTACAGAGAAAGCAATCTGTGTGGCTAGTGGGCAGATTACCATGCAAGCCCCAGGAGA  
AATGGAGGAGCTTTGTAGCCACCTCCCTGTCAGCCAGTATTAACATGTCCCTTCCCCCT  
GCCCCGCCGTAGATTACAGGACATTGCCCCCTGTGTGCCACCAAACCAGGACTTCCCCCT  
GGCTTGGCATCCCTGGCTCTCTCCTGGTACCCAGCAAGACGTCTGTTCCAGGGCAGTGTA  
GCATCTTTCAAGCTCCGTTACTATGGCGATGGCCATGATGTTACAATCCCACTTGCCTGA  
ATAATCAAGTGGGAAGGGGAAGCAGAGGGAAATGGGGCCATGTGAATGCAGCTGCTCTGT  
TCTCCCTACCCCTGAGGAAAAACCAAAGGGAAGCAACAGGAACCTTCTGCAACTGGTTTTTA  
TCGGAAGATCATCCTGCCTGCAGATGCTGTTGAAGGGGCACAAGAAATGTAGCTGGAGA  
AGATTGATGAAAGTGCAGGTGTGTAAGGAAATAGAACAGTCTGCTGGGAGTCAGACCTGG  
AATTCTGATTCCAACTCTTTATTACTTTGGGAAGTCACTCAGCCTCCCCGTAGCCATCT  
CCAGGGTGACGGAACCCAGTGTATTACCTGCTGGAACCAAGGAACTAACAATGTAGGTT  
ACTAGTGAATACCCCAATGGTTTCTCCAATTATGCCCATGCCACCAAACAATAAAACAA  
AATTCTCTAACACTGAAA

160/392

## FIGURE 154

MWLPLGLLSLCLSPILSSPSLKSQACQQLWTLPSPLVAFRANRTTYVMDVSTNQSG  
MEHRNHLCFCDLYDRATSPPLKCSLL



**FIGURE 155**

GTAGACGCGTCTTGGGTCTCCCGGCTGCCGCTGCTGCCGCCGCCGCCCTCGGGTCGTGGAGC  
CAGGAGCGACGTCAACGCCATGGCAGGCATCAAAGCTTTGATTAGTTTGTCTTTGGAGG  
AGCAATCGGACTGATGTTTTTGATGCTTGGATGTGCCCTTCCAATATACAACAAATACTG  
GCCCCCTCTTTGTTCTATTTTTTTTACATCCTTTCACCTATTCCATACTGCATAGCAAGAAG  
ATTAGTGGATGATACAGATGCTATGAGTAACGCTTGTAAGGAACCTGCCATCTTTCTTAC  
AACGGGCATTGTGTCGTGTCAGCTTTTGGACTCCCTATTGTATTTGCCAGAGCACATCTGAT  
TGAGTGGGGAGCTTGTGCACCTGTTCTCACAGGAAACACAGTCATCTTTGCAACTATAC  
AGGCTTTTTCTTGGTCTTTGGAAGCAATGACGACTTCAGCTGGCAGCAGTGGTGA<sup>1</sup>AAAAG  
AATTACTGAACTATTGTCAAATGGACTTCCTGTCAATTTGTTGGCCATTACGCACACAGG  
AGATGGGGCAGTTAATGCTGAATGGTATAGCAAGCCTCTTGGGGGTATTTTAGGTGCTCC  
CTTCTCAC<sup>2</sup>TTTTATTGTAAGCATACTATTTTCACAGAGACTTGCTGAAGGATTA<sup>3</sup>AAAGGA  
TTTTCTCTTTTGGAAAAGCTTGACTGATTTTCACTTATCTATAGTATGCTTTTTTGTGGT  
GTCCTGCTGAATTTAAATATTTATGTGTTTTTCTGTTAGGTTGATTTTTTTTGGAAATCA  
ATATGCAATGTTAAACACTTTTTTAAATGTAATCATTTGCATTGGTTAGGAATTCAGAATT  
CCGCCGGCTCTATTACTGGTCAAGTACATCTTTTCTCTTAA<sup>4</sup>AATTATTTAGCCTCCATTA  
TTACAAAAAATTATAAAAATAAGTTTTTCAGTCAGTCAGGATGACATCACTCCCAATGTTA  
TGCAGACATACAGACGGTTGGCATAACGTTATAGACTGTATACTCAGTGCAAATATAGCTG  
CATTTATACCTCAGAGGGGCCAAGTGTTAATGCCCATGCCCTCCGTTAAGGGTTGTTGGT  
TTTACTGTTAGACAGATGTTTTTGTGGAATTGAAAATTATTTTATGGAATTGCTACAGAGGA  
GTGCTTTTCTTCTCAATTGTTTAGAAGAATTATGTTTAAACTTTAAGGTAAAGGGTGTAAAA  
ACATTTT<sup>5</sup>TGAGATAAGGTTTTTATTTATGTTTATTGTTATGTTAGAGTGAGTTGCAATGTGG  
GAAGAAATGACATTGAAATTCAGTTTTTGAATCCTGTTTTCTATTTATAAGTGAAATTTG  
TGATCTCCTATCAACCTTTCATGTTTTACCCTGTTAA<sup>6</sup>AATGGACATACATGGAACCACTA  
CTGATGAGGGACAGTTGTATGTTTGCATCATATATGCCAGAAAACCTTCCTCTGCTTCCT  
CTTTTGACTTATTTGGTATGTTGTATATATTACATAAAATAACTTTTTCAAATATAGTTT  
AATAACACTTAGAAGTGTTTACTTACCTGGAAAATAATTGCTATGCCGTACATT<sup>7</sup>CAGAGT  
GCCCCCTCCCCTGCAAGGCCTTGCCATGATTAACAAGTAACTTGTTAGTCTTACAGATAA  
TTCATGCATTAAACAGTTTAAAGATTTAGACCATGGTAATAGTAGTTCTTATTCTCTAAGGT  
TATATCATATGTAATTTAAAAGTATTTTAAAGACAAGTTTCTGTATACCTCTGAACTGT  
TTTGATTTTGAGTTCATCATGATAGATCTGCTGTTTCCTTATAAAAGGCATTTGTTGTGT  
GAGTTAATGCAAAGTAGCCAAGTCCAGCTATATAGCAGCTTCAGAAACATACCTGACCAA  
AAAATTCCCAGTAACCAGGCATGATCAATTTATAGTGGT<sup>8</sup>CGTTTACATCTAATAATTATC  
AGGACTTTTTTTCAGGAGTGGGTTATAAAAAACATTCAGTTGGTCTGACAGTATTTTGTTA  
AGGATATTTGTTTGTATGTTTATT<sup>9</sup>CAGTATACTTACATAAAAATTATTTTCGCCATCAGCC  
AAAAC<sup>10</sup>TCAGTAATCATGACAGCTGTCTGTTGTTTTATGAAGTTTATTTCTCAAGAAAATG  
GGAATAAAATTTGGGATTTGTTTCAGCTTTTTTACTAAAGATGCCTAAAGCCACAGGTTTTA  
TTGCCTAACTTAAGCCATGACTTTTAGATATGAGATGACGGGAAGCAGGACGAAATATCG  
GCGTGTGGCTGGAGCCTTCCCACTGGAGGCTGAAAGTGGCTTGTGGTATTATAATGTTCA  
GATTTCAAGAGGAAGGTGCAGGTACACATGAGTTAGAGAGCTGGTGAGACAGTTGGGAAC  
TCTTTGTGCTTGTGATCTACTGGACTTTTTTTTTCAGGAAGTGCATTCTCTGGTCTTTC  
CCTATTTTCTGTTCTGGATGTCACTGGAGTGCAGTGCAGTCTACTGTTTTATCCACTTGGCCAC  
AGACTTTTTTCTAACAGCTGCGTATTATTTCTATATACTAATTGATTCATGGCAGCATTGTGT  
CTTTGACCTTGTATACTAGCTTGACATAGTGTCTGTCTGATTGTTCTAGGCTAGTTACTGTG  
AGATATGAATTTTCCATAGAATATGCAGTGATACAACATTACCATTCTTCTATGGAAAGA  
AAACTTTTTGATGATGAAACAATAAAGATTTTAAATATCTATTTTAAAAAAAAAAAA

162/392

## FIGURE 156

MAGIKALISLSFGGAIGLMFLMLGICALPIYNKYWPLFVLFFYILSPIPYCIARRLVDDTD  
AMSNACKELAIFLTGTGIVVSAFGLPIVFARAHLEWGACALVLTGNTVIFATILGFFLVF  
GSNDDFSQQW

163/392

## FIGURE 157

GTTTCTCATAGTTGGCGTCTTCTAAAGGAAAAACACTAAAATGAGGAACTCAGCGGACCG  
GGAGCGACGCAGCTTGAGGGAAGCATCCCTAGCTGTTGGCGCAGAGGGGCGAGGCTGAAG  
CCGAGTGGCCCGAGGTGTCTGAGGGGCTGGGGCAAAGGTGAAAGAGTTTCAGAACAAAGCT  
TCCTGGAACCCATGACCCATGAAGTCTTGTCGACATTTATACCGTCTGAGGGTAGCAGCT  
CGAAACTAGAAGAAGTGGAGTGTGTCAGGGACGGCAGTATCTCTTTGTGTGACCCTGGC  
GGCCTATGGGACGTTGGCTTCAGACCTTTGTGATACACCATGCTGCGTGGGACGATGACG  
GCGTGGAGAGGAATGAGGCCTGAGGTACACTGGCTTGCCCTCCTAGCCACAGCAGGC  
TGCTTTGCTGACTTGAACGAGGTCCCTCAGGTCACCGTCCAGCCTGCGTCCACCGTCCAG  
AAGCCCGGAGGCACTGTGATCTTGGGCTGCGTGGTGGAACTCCAAGGATGAATGTAACC  
TGGCGCCTGAATGGAAGGAGCTGAATGGCTCGGATGATGCTCTGGGTGTCTCATCACC  
CACGGGACCCTCGTCATCACTGCCCTTAACAACCACTGTGGGACGGTACCAGTGTGTG  
GCCCCGATGCTGCGGGGGCTGTGGCCAGCGTGCCAGCCACTGTGACACTAGCCAATCTC  
CAGGACTTCAAGTTAGATGTGCAGCAGTGATGAAAGTGGATGAGGGAAACACAGCAGTC  
ATTGCTTGCACCTGCCTGAGAGCCACCCCAAAGCCAGGTCCGGTACAGCGTCAAACAA  
GAGTGGCTGGAGGCCTCCAGAGGTAACCTGATCATGCCCTCAGGGAACCTCCAGATT  
GTGAATGCCAGCCAGGAGGACGAGGGCATGTACAAGTGTGCAGCCTACAACCCAGTGACC  
CAGGAAGTGAAAACCTCCGGCTCCAGCGACAGGCTACGTGTGCGCCGCTCCACCGCTGAG  
GCTGCCCGCATCATCTACCCCCCAGAGGCCCAAACCATCATCGTCACCAAAGGCCAGAGT  
CTCATTTCTGGAGTGTGTGGCCAGTGGAATCCACCCCCACGGGTACCTGGGCCAAGGAT  
GGGTCCAGTGTACCGGCTACAACAAGGCTTGGCAGGATGCTGAGCTGGCTACTGGCACA  
ACCACCAGCGAGGAGGACTCAGGCACCTACCGCTGCATGGCCGACAATGGGGTTGGGCAG  
CCCCGGGCGAGCGGTCTCTCTACAATGTCCAGGTGTTTGAACCCCCCTGAGGTCAACATG  
GAGCTATCCCAGCTGGTCTATCCCTTGGGGCCAGAGTGCCAAGCTTACCTGTGAGGTGCGT  
GGGAACCCCCCGCCCTCCGTGCTGTGGCTGAGGAATGCTGTGCCCTCATCTCCAGCCAG  
CGCCTCCGGCTCTCCCGCAGGGCCCTGCGCGTGCTCAGCATGGGGCCTGAGGACGAAGGC  
GTCTACAGTGCATGGCCGAGAACGAGGTTGGGAGCGCCCATGCCGTAGTCCAGCTGCGG  
ACCTCCAGGCCAAGCATAACCCCAAGGCTTGGCAGGATGCTGAGCTGGCTACTGGCACA  
CCTCCTGTATCACCTTCAAACCTCGGCAACCTTGAGCAGATGCTGAGGGGGCAACCGGCG  
CTCCCCAGACCCCCAACGTCACTGGGGCCTGCTTCCCCGAAGTGTCCAGGAGAGAAGGGG  
CAGGGGGCTCCCCCGGAGGCTCCCATCATCTCAGCTCGCCCCGCACCTCCAAGACAGAC  
TCATATGAACTGGTGTGGCGGCCTCGGCATGAGGGCAGTGGCCGGGCGCCAATCCTCTAC  
TATGTGGTGAAACACCGCAAGCAGGTCAAAATTCCTCTGACGATTGGACCATCTCTGGC  
ATTTCAGACCAACAGCACCGCCTGACCTTACCAGACTTGACCCCGGAGCTTGTATGAA  
GTGGAGATGGCAGCTTACAACCTGTGCGGGAGAGGGCCAGACAGCCATGGTCACTTCCGA  
ACTGGACGGCGGCCCAAACCCGAGATCATGGCCAGCAAAGAGCAGCAGATCCAGAGAGAC  
GACCCTGGAGCCAGTCCCCAGAGCAGCAGCCAGCCAGACCACGGCCGCTCTCCCCCCCCA  
GAAGCTCCCGACAGGCCACCATCTCCACGGCCTCCGAGACCTCAGTGACCTGACCTGG  
ATTCCCCGTGGGAATGGTGGGTTCCTCAATCCAGTCTTCCGTGTGGAGTACAAGAAGCTA  
AAGAAAGTGGGAGACTGGATTCTGGCCACCAGCGCCATCCCCCATCGCGGCTGTCCGTG  
GAGATCACGGGCCTAGAGAAAGGCACCTCTACAAGTTTCGAGTCCGGGCTCTGAACATG  
CTGGGGGAGAGCGAGCCAGCGCCCCCTCTCGGCCCTACGTGGTGTGCGGGCTACAGCGGT  
CGCGTGTACGAGAGGCCCCGTGGCAGGTCTTATATCACCTTACGGATGCGGGTCAATGAG  
ACCACCATCATGCTCAAGTGGATGTACATCCAGCAAGTAACAACAACACCCCAATCCAT  
GGCTTTTATATCTATTATCGACCCACAGACAGTGACAATGATAGTGAATAACAAGGAT  
ATGGTGGAAAGGGGACAAGTACTGGCACTCCATCAGCCACCTGCAGCCAGAGACCTCCTAC  
GACATTAAGATGCAGTGCTTCAATGAAGGAGGGGAGAGCGAGTTTACGCAACGTGATGATC  
TGTGAGACCAAAGCTCGGAAGTCTTCTGGCCAGCCTGGTTCGACTGCCACCTCCAACTCTG  
GCCCCACCACAGCGCCCTTCTGAAACCATAGAGCGGCGGTGGGCATGGGGCCATG  
GTGGCTCGCTCCAGCGACCTGCCCTATCTGATTGTGCGGGTCTGCTGGGCTCCATCGTT  
CTCATCATCGTCACCTTCATCCCCTTCTGCTTGTGGAGGGCCTGGTCTAAGCAAAAACAT  
ACAACAGACCTGGGTTTTCTCGAAGTGCCCTTCCACCTCCTGCCCCGTATACTATGGTG  
CCATTGGGAGGACTCCAGGCCACCAGGCCAGTGACAGCCCTACCTCAGTGGCATCAGT

164/392

GGACGGGCCTGTGCTAATGGGATCCACATGAATAGGGGCTGCCCCCTCGGCTGCAGTGGGC  
TACCCGGGCATGAAGCCCCAGCAGCACTGCCCAGGCGAGCTTCAGCAGCAGAGTGACACC  
AGCAGCCTGTGCTGAGGCAGACCCATCTTGGCAATGGATATGACCCCCAAAGTCACCAGATC  
ACGAGGGGTCCCAAGTCTAGCCCGGACGAGGGCTCTTTCTTATACACACTGCCCCGACGAC  
TCCACTCACCAGCTGCTGCAGCCCCATCACGACTGCTGCCAACGCCAGGAGCAGCCTGCT  
GCTGTGGGCCAGTCAGGGGTGAGGAGAGCCCCGACAGTCCTGTCTGGAAGCAGTGTGG  
GACCCCTCCATTTCACTCAGGGCCCCCATGCTGCTTGGGCCTTGTGCCAGTTGAAGAGGTG  
GACAGTCCTGACTCCTGCCAAGTGAGTGGAGGAGACTGGTGTCCCCAGCACCCCGTAGGG  
GCCTACGTAGGACAGGAACCTGGAATGCAGCTCTCCCCGGGGCCACTGGTGCGTGTGTCT  
TTTGAAACACCACCTCTCACAATTTAGGCAGAAGCTGATATCCAGAAAGACTATATATT  
GTTTTTTTTTTAAAAAAAAGAGAGAAAAAGAGACAGAGAAAATTGGTATTTATTTTC  
TATTATAGCCATATTTATATATTTATGCACTTGTAATAAATGTATATGTTTTATAATTC  
TGGAGAGACATAAGGAGTCCTACCCGTTGAGGTGAGAGAGGGAAAATAAAGAAGCTGCCA  
CCTAACAGGAGTCACCCAGGAAAGCACCGCACAGGCTGGCGCGGGACAGACTCCTAACCT  
GGGGCCTCTGCAGTGGCAGGCGAGGCTGCAGGAGGCCACAGATAAGCTGGCAAGAGGAA  
GGATCCCAGGCACATGGTTCATCACGAGCATGAGGGAACAGCAAGGGGCACGGTATCACA  
GCCTGGAGACACCCACACAGATGGCTGGATCCGGTGCTACGGGAAACATTTTCCTAAGAT  
GCCCATGAGAACAGACCAAGATGTGTACAGCACTATGAGCATTAAAAACCTTCCAGAAT  
CAATAATCCGTGGCAACATATCTCTGTAAAAACAACACTGTAACCTCTAAATAAATGTT  
TAGTCTTCCCTGTAAAA

165/392

**FIGURE 158**

MLRGTMATAWRGMRPEVTLACLILLATAGCFADLNEVPQVTVQPASTVQKPGGTIVILGCVVE  
PPRMNVTWRLNGKELNGSDDALGVLITHGTLVITALNNHTVGRYQCVARMPAGAVASVPA  
TVTLANLQDFKLDVQHVIEVDEGNTAVIACHLPESHKPAQVRYSVKQEWLEASRGNYLIM  
PSGNLQIVNASQEDEGMYKCAAYNPVTQEVKTS GSSDRLRVRRSTAEAAARI IYPPEAQTI  
IVTKGQSLILECVASGI PPPRVTWAKDGSSVTGYNKTRFLLSNLLIDTTSEEDSGTYRCM  
ADNGVGQPGAAVILYNVQVFEPPEVTMELS QLVIPWGQSAKLTCEVRGNPPPSVLWLRNA  
VPLISSQRLRLSRRLRVLSMGPEDEGVYQCMANEVGS AHAVVQLRTSRPSITPRLWQD  
AELATGTTPPVSPSKLGNPEQMLRGQPALPRPPTS VGPASPKCPGEGKGQGAPEAPIILSS  
PRTSKTDSYELVWRPRHEGSGRAPILYYVVKHRKQVTNSSDDWTISGIPANQHRLTLTRL  
DPGSLYEVEMAAYNCAGEGQTAMVTFRTGRRPKPEIMASKEQQIQRDDPGASPOSSSQPD  
HGR LSPPEAPDRPTISTASETSVYVTWIPRGNGGFPIQSFRVEYKKLKKVGDWILATSAI  
PPSRLSVEITGLEKGT SYKFRVRALNMLGESEPSAPSRPYVVS GYSGRVYERPVAGPYIT  
FTDAVNETTIMLKWMIYPASNNNTPIHGFIYYRPTSDNDSYK KDMVEGDKYWHSISH  
LQPETSYDIKMQCFNEGGESEFSNVMICETKARKSSGQPGR LPPPTLAPPQPLPETIER  
PVG TGAMVARSSDLPYLIVGVVLGSIVLII VTFIPFCLWRAW SKQKHTTDLGFPR SALPP  
SCPYTMVPLGGLPGHQASGQPYLSGISGRACANGIHMNRGCPSAAVGYPGMKPQQHCPGE  
LQQQSDTSSLLRQTHLGNGYDPQSHQITRGPKSSPDEGSFLYTL PDDSTHQLLQPHHDC  
QRQEQPAAVGQSGVRRAPDSPVLEAVWDPPFHS GPPCCLGLVPVEEVDSPDSCQVSGGDW  
CPQHVPV GAYVGQEPGMQLSPGPLVRVVSFETPPLTI

**Signal peptide:**

amino acids 1-30

**Transmembrane domain:**

amino acids 16-30 (type II), 854-879

166/392

**FIGURE 159**

CCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCG  
CCCACGCGTCCGCCCACGCGTCCGGTGCAAGCTCGCGCCGCACACTGCCTGGTGAGGGA  
AGGAGCCCCGGGCGCCTCTCGCCGCTCCCCGCGCCGCGTCCGCACCTCCCCACCGCCCGC  
CGCCCGCCGCCCCGCGCCCGCAAAGCATGAGTGAGCCCGCTCTCTGCAGCTGCCCGGGC  
GCGAATGGCAGGCTGTTTCCGCGGAGTAAAAGGTGGCGCCGGTCAGTGGTCGTTTCCAAT  
GACGGACATTAACCAGACTGTCAGATCCTGGGGAGTCGCGAGCCCCGAGTTTGGAGTTTT  
TTCCCCCACAACGTCACAGTCCGAAGTGCAGAGGGAAGGAAGGCGGCAGGAAGGCGAA  
GCTCGGGCTCCGGCACGTAGTTGGGAAACTTGCGGGTCTAGAAAGTCGCCTCCCCGCCTT  
GCCGCGCGCCCTTGCAGCCCCGAGCCGAGCAGCAAAGTGAGACATTGTGCGCCTGCCAGA  
TCCGCGCGCCGCGGACCGGGGCTGCCTCGGAAACACAGAGGGGTCTTCTCTCGCCCTGCA  
TATAATTAGCCTGCACACAAAGGGAGCAGCTGAATGGAGGTTGTCACTCTCTGGAAAAGG  
ATTTCTGACCGAGCGCTTCCAATGGACATTCTCCAGTCTCTCTGGAAAGATTCTCGCTAA  
TGGATTTCTGCTGCTCGGTCTCTGTCTATACTGGCTGCTGAGGAGGCCCTCGGGGGTGG  
TCTTGTGTCTGCTGGGGCCTGCTTTCAGATGCTGCCCGCCGCCCCCAGCGGGTGCCCGC  
AGCTGTGCCGGTGCGAGGGGCGGCTGCTGTACTGCGAGGCGCTCAACCTCACCGAGGCGC  
CCCACAACCTGTCCGGCCTGCTGGGCTTGTCCCTGCGCTACAACAGCCTCTCGGAGCTGC  
GCGCCGCGCAGTTCACGGGGTTAATGCAGCTCACGTGGCTCTATCTGGATCACAATCACA  
TCTGCTCCGTGCAGGGGGACGCCTTTCAGAACTGCGCCGAGTTAAGGAACTCACGCTGA  
GTTCCAACCAGATCACCCAACCTGCCCAACACCACTTCCGGCCCCATGCCCAACCTGCGCA  
GCGTGGACCTCTCGTACAACAAGCTGCGAGGCGCTCGCGCCCGACCTCTTCCACGGGCTGC  
GGAAGCTCACCAAGCTGCATATGCGGGCCAACGCCATCCAGTTTGTGCCCCGTGCGCATCT  
TCCAGGACTGCCGCAGCCTCAAGTTTCTCGACATCGGATAACAATCAGCTCAAGAGTCTGG  
CGCGCAACTCTTTCGCCGGCTTGTTTAAGCTCACCGAGCTGCACCTCGAGCACAACGACT  
TGGTCAAGGTGAACTTCGCCCACTTCCCGCGCCTCATCTCCCTGCACTCGCTCTGCCTGC  
GGAGGAACAAGGTGGCCATTGTGGTCAGCTCGCTGGACTGGGTTTGGAACTGGAGAAAA  
TGGACTTGTGCGGCAACGAGATCGAGTACATGGAGCCCCATGTGTTGAGACCGTGCCGC  
ACCTGCAGTCCCTGCAGCTGGACTCCAACCGCCTCACCTACATCGAGCCCCGGATCCTCA  
ACTCTTGGAAGTCCCTGACAAGCATCACCTGGCCGGGAACCTGTGGGATTGCGGGCGCA  
ACGTGTGTGCCCTAGCCTCGTGGCTCAGCAACTTCCAGGGGCGCTACGATGGCAACTTGC  
AGTGCGCCAGCCCGGAGTACGCACAGGGCGAGGACGTCCTGGACGCCGTGTACGCCCTTCC  
ACCTGTGCGAGGATGGGGCCGAGCCACAGCGGCCACCTGCTCTCGGCCGTACCAACC  
GCAGTGATCTGGGGCCCCCTGCCAGCTCGGCCACCAAGCTCGCGGACGGCGGGGAGGGG  
AGCACGACGGCACATTCGAGCCTGCCACCGTGGCTCTTCCAGGCGGCGAGCACGCCGAGA  
ACGCCGTGCAGATCCACAAGGTGGTCACGGGCACCATGGCCCTCATCTTCTCCTTCTCA  
TCGTGGTCTTGGTGCTCTACGTGTCTTGGAAAGTGTTCAGCCAGCCTCAGGCAGCTCA  
GACAGTGCTTTGTACGCAGCGCAGGAAGCAAAAGCAGAAACAGACCATGCATCAGATGG  
CTGCCATGTCTGCCAGGAATACTACGTTGATTACAAACCGAACCACATTGAGGGAGCCC  
TGGTGATCATCAACGAGTATGGCTCGTGACCTGCCACCAGCAGCCCGCGAGGGAATGCG  
AGGTGTGATTGTCCAGTGGCTCTCAACCATGCGCTACCAAATACGCCTGGGCAGCCGG  
GACGGGCCGCGGGCACCAGGCTGGGGTCTCCTTGTCTGTGCTCTGATATGCTCCTTGAC  
TGAAACTTTAAGGGGATCTCTCCAGAGACTTGACATTTTAGCTTTATTGTGTCTTAAAC  
ACAAAAGCGAATTAAAAACACAACAAAAACCCACCCACCAACCTTCAGGACAGTCTAATC  
TTAAATTTTCATATGAGAACTCCTTCTCCCTTTGAAGATCTGTCCATATTAGGAATCTG  
AGAGTGTAATAAAGGTGGCCATAAGACAGAGAGAGAATAATCGTGCTTTGTTTTATGCTA  
CTCCTCCACCCCTGCCCATGATTAAACATCATGTATGTAGAAGATCTTAAGTCCATACGC  
ATTTTCATGAAGAACCATTGGAAAGAGGAATCTGCAATCTGGGAGCTTAAGAGCAAATGAT

167/392

GACCATAGAAAGCTATGTTCTTACTTTGTGTGTGTGTCTGTATGTTTCTGCGTTGTGTGT  
CTTTGTAGGCAAGCAAACGTTGTCTACACAAACGGGAATTTAGCTCACATCATTTTCATGC  
CCCTGTGCCTCTAGCTCTGGAGATTGGTGGGGGAGGTGGGGGGAAACGGCAGGAATAAG  
GGAAAGTGGTAGTTTTAACTAAGGTTTTGTAACTTGAATCTTTTCTTCTCAAATTA  
ATTATCTTTAAGCTTCAAGAACTTGCTCTGACCCCTCTAAGCAAATACTAAGCATTTA  
AAAGAGAATCTAATTTTTAAAGGTGTAGCACCTTTTTTTTTTATTCTTCCCACAGAGGGTG  
CTAATCTCATTATGCTGTGCTATCTGAAAAGAACTTAAGGCCACAATTCACGTCTCGTCC  
TGGGCATTGTGATGGATTGACCCTCCATTTGCAGTACCTTCCCAGCTGATTAAAGTTCAG  
CAGTGGTATTGAGGTTTTTTCGAATATTTATATAGAAAAAAGTCTTTTTCACATGACAAAT  
GACACTCTCACACCAGTCTTAGCCCTAGTAGTTTTTTTAGGTTGGACCAGAGGAAGCAGGT  
TAAATGAGACCTGTCTCTGCTGCACTCAGAAAAAATAGGCAGTCCCTGATGCTCAGATC  
TTAGCCTTGATATTAATAGTTGAGACCACCTACCCACAATGCAGCCTATACTCCCAAGAC  
TACAAAGTTACCATCGCAAAGGAAAGGTTATTCCAGTAAAAGGAAATAGTTTTCTCAACC  
ATTTAAAAATATTCTTCTGAACTCATCAAAGTAGAAGAGCCCCAACCTTTTCTCTCTGC  
CTTCAAGAAGGCAGACATTTGGTATGATTTAGCATCAACAACACATTTATGAGTATATGT  
AAGTAATCAGAGGGGCAAATGCCACTTGTTATTCCTCCCAAGTTTCCAAGCAAGTACAC  
ACAGATCTCTGGTAGGATTAGGGGCCACTTGTGTTTCCGGCTTATTTTAGTCGACTTGTC  
AGCAAGTTTGATGCCTAGTCTATCTGACATGGCCAGTAGAACAGGGCATTGATGGATCA  
CATGAGATGGTAGAAGGAACATCATCACATACCCCTCTCACAGAGAAAATTATCAAAGAA  
CCAGAAATTATATCTGTTTTGGAGCAAGAGTGTCATAATGTTTCAGGGTAGTCAAATAA  
ACATAAATTATCTCTCTAGATGAGTGGCGATGTTGGCTGATTTGGGTCTGCCATTGACA  
GAATGTCAAATAAAAAGGAATTAGCTAGAATATGACCATTAAATGTGCTTCTGAAATATA  
TTTTGAGATAGGTTTAGAATGTCA

168/392

**FIGURE 160**

MDFLLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEA  
PHNLSGLLGLSLRYNSLSELRAQQFTGLMQLTWLYLDHNNHICSVQGDAFQKLRRVKELTL  
SSNQITQLPNTTFRPMPNLRSDLSYNKLQALAPDLFHGLRKLTTLHMRANAIQFVPVRI  
FQDCRSKFLDIGYNQLKSLARNSFAGLFKLTELHLEHNDLVKVNFAHFPRILISLHSLCL  
RRNKVAIVVSSLDWVWNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDSNRLTYIEPRIL  
NSWKSLSITLAGNLWDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDVAVAF  
HLCEDGAEPTSGHLLSAVTNRSDLGPPASSATTADGGEGQHDGTFEPATVALPGGEHAE  
NAVQIHKVVTGTMALIFSFLIVVLVLYVSWKCFPASLRQLRQCFVTQRRKQKQKQTMHQM  
AAMSAQEYYVDYKPNHIEGALVIINEYGSCTCHQQPARECEV





170/392

**FIGURE 162**

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPESSTKETERKETKAE  
ELDAEVLEVFPHTHEWQALQPGQAVPAGSHVRLNLQTGEREAKLOYEDKFRNNLKGKRLD  
INTNTYTSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIET  
DMQIMVRLINKFNSSSSSLEEKIAALFDLEYVYVHQMDNAQDLLSFGLQVINGLNSTEP  
LVKEYAAAFVLGAASFSSNPKVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHF  
PYAQRQFLKLGGLOVLRRTLVOEKGTEVLAVRVVTLVLDLVTEKMFEEEEAELTQEMSPEK  
LQQYRQVHLLPGLWEQGWCEITAHLLALPEHDAREKVLQTLGVLLTTCRDRYRQDPQLGR  
TLASLQAEYQVLASLELQDGEDEGYFQELLGSVNSLLKELR

**Important features:**

**Signal peptide:**

amino acids 1-29

**Hypothetical YJL126w/YLR351c/yhcX family protein:**

amino acids 364-373

**N-glycosylation site:**

amino acids 193-197, 236-240

**N-myristoylation site:**

amino acids 15-21, 19-25, 234-240, 251-257, 402-408, 451-457

**Homologous region SLS1 protein:**

amino acids 68-340

171/392

**FIGURE 163**

CAGAGAGGAGGCTTTGGGAATTGTCCAGCAGAAACAGAGAAGTCTGAGGTGGTGTCAAGA  
CAAAAGATGCTTCAGCTTTGGAACTTGTTCTCCTGTGCGGCGTGCTCACTGGGACCTCA  
GAGTCTCTTCTTGACAATCTTGGAATGACCTAAGCAATGTCTGGATAAGCTGGAACCT  
GTTCTTCACGAGGGACTTGAGACAGTTGACAATACTCTTAAAGGCATCCTTGAGAACTG  
AAGGTCGACCTAGGAGTGCTTCAGAAATCCAGTGCTTGGCAACTGGCCAAGCAGAAGGCC  
CAGGAAGCTGAGAAATTGCTGAACAATGTCATTTCTAAGCTGCTTCCAATAACACGGAC  
ATTTTTGGGTGAAAATCAGCAACTCCCTCATCCTGGATGTCAAAGCTGAACCGATCGAT  
GATGGCAAAGGCCTTAACCTGAGCTTCCCTGTCACCGCGAATGTCACTGTGGCCGGGCCC  
ATCATTGGCCAGATTATCAACCTGAAAGCCTCCTTGGACCTCCTGACCGCAGTCACAATT  
GAAACTGATCCCCAGACACACCAGCCTGTTGCCGTCTTGGGAGAATGCGCCAGTGACCCA  
ACCAGCATCTCACTTTCCTTGCTGGACAAACACAGCCAAATCATCAACAAGTTCGTGAAT  
AGCGTGATCAACACGCTGAAAAGCACTGTATCCTCCCTGCTGCAGAAGGAGATATGTCCA  
CTGATCCGCATCTTCATCCACTCCCTGGATGTGAATGTCATTGAGCAGGTCGTCGATAAT  
CCTCAGCACAAAACCCAGCTGCAAACCCTCATCTGAAGAGGACGAATGAGGAGGACCACT  
GTGGTGCATGCTGATTGGTTCCAGTGGCTTGCCCCACCCCCTTATAGCATCTCCCTCCA  
GGAAGCTGCTGCCACCACCTAACCAGCGTGAAAGCCTGAGTCCCACCAGAAGGACCTTCC  
CAGATACCCCTTCTCCTCACAGTCAGAACAGCAGCCTCTACACATGTTGTCCTGCCCTG  
GCAATAAAGGCCCATTTCTGCACCTTAA

172/392

## FIGURE 164

MLQLWKLVLLCGVLTGTSESLLDNLGNDLSNVVDKLEPVLHEGLETVDNTLKGILEKLKV  
DLGVLQKSSAWQLAKQKAQEA EKLLNNVISKLLPTNTDIFGLKISNSLILDVKAEPIDDG  
KGLNLSFPVTANVTVAGPIIGQIINLKASLDLLTAVTIETDPQTHQPVAVLGECASDPTS  
ISLSLLDKHSQIINKFVNSVINTLKSTVSSLLQKEICPLIRIFIHSLDVNVIQQVVDNPQ  
HKTQLQTLI

### Important features:

#### Signal peptide:

1-15

#### Transmembrane domain:

none

#### N-glycosylation site:

124-128, 132-136

#### N-myristoylation site:

12-18, 16-22, 26-32, 101-107, 122-128, 141-147

#### Leucine zipper pattern:

44-66

173/392

**FIGURE 165**

GCAGTCAGAGACTTCCCCTGCCCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCT  
CTTGCTTCCTGAACTAGCTCACAGTAGCCCGGCGGCCAGGGCAATCCGACCACATTTCA  
CTCTCACCGCTGTAGGAATCCAGATGCAGGCCAAGTACAGCAGCACGAGGGACATGCTGG  
ATGATGATGGGGACACCACCATGAGCCTGCATTCTCAAGCCTCTGCCACAACCTCGGCATC  
CAGAGCCCCGGCGCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGA  
CCCTGCTGACTTTGTGCTTGGTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGT  
TTCAGTACTACCAGCTCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGAT  
TAGGAAATACGTCCCAAGAGTTGCAATCTCTTCAAGTCCAGAATATAAAGCTTGCAGGAA  
GTCTGCAGCATGTGGCTGAAAACTCTGTGCTGAGCTGTATAACAAAGCTGGAGCACACA  
GGTGCAGCCCTTGTACAGAACAATGGAAATGGCATGGAGACAATTGCTACCAGTTCTATA  
AAGACAGCAAAAGTTGGGAGGACTGTAAATATTTCTGCCTTAGTGAAAACCTACCATGC  
TGAAGATAAAACAAGAAGACCTGGAATTTGCCGCGTCTCAGAGCTACTCTGAGTTTT  
TCTACTCTTATTGGACAGGGCTTTTGCGCCCTGACAGTGGCAAGGCCTGGCTGTGGATGG  
ATGGAACCCCTTTCACCTCTGAACTGTTCCATATTATAATAGATGTCACCAGCCCAAGAA  
GCAGAGACTGTGTGGCCATCCTCAATGGGATGATCTTCTCAAAGGACTGCAAAGAATTGA  
AGCGTTGTGTCTGTGAGAGAAGGGCAGGAATGGTGAAGCCAGAGAGCCTCCATGTCCCCC  
CTGAAACATTAGGCGAAGGTGACTGATTCGCCCTCTGCAACTACAAATAGCAGAGTGAGC  
CAGGCGGTGCCAAAGCAAGGGCTAGTTGAGACATTGGGAAATGGAAATAATCAGGAAAG  
ACTATCTCTCTGACTAGTACAAAATGGGTCTCGTGTTTCTGTTCAGGATCACCAGCAT  
TTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACAAGAAGTCTTATTTACATGCCAC  
CAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTGGCTTAGAGATAACTTTTAGC  
TCTCTTTCTTCTCAATGTCTAATATCACCTCCCTGTTTTCATGTCTTCCTTACACTGGT  
GGAATAAGAACTTTTTGAAGTAGAGGAAATACATTGAGGTAACATCCTTTTCTCTGACA  
GTCAAGTAGTCCATCAGAAATTGGCAGTCACTTCCCAGATTGTACCAGCAAATACACAAG  
GAATTCTTTTTGTTTGTTCAGTTCATACTAGTCCCTTCCCAATCCATCAGTAAAGACCC  
CATCTGCCTTGTCCATGCCGTTTTCCCAACAGGGATGTCACCTTGATATGAGAATCTCAAAT  
CTCAATGCCTTATAAGCATTCCTTCTGTGTCCATTAAAGACTCTGATAATTGTCTCCCT  
CCATAGGAATTTCTCCAGGAAAGAAATATATCCCCATCTCCGTTTCATATCAGAACTAC  
CGTCCCCGATATTCCCTTCCAGAGAGATTAAAGACCAGAAAAAAGTGAGCCTCTTCATCTG  
CACCTGTAATAGTTTCAGTTCCTATTTTCTTCCATTGACCCATATTTATACCTTTCAGGT  
ACTGAAGATTTAATAATAATAAATGTAAATACTGTGAAAAA

174/392

**FIGURE 166**

MQAKYSSTRDMLDDDGD'TTMSLHSQASATTRHPEPRRTEHRAPSSTWRPVALTLLTLCLV  
LLIGLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEK  
LCRELYNKAGAHRCSPCTEQWKWHGDNCYQFYKDSKSWEDCKYFCLSENSTMLKINKQED  
LEFAASQSYSEFFYSYWTGLLRPDSGKAWLWMDGTPFTSELFHIIIDVTSPRSRDCVAIL  
NGMIFSKDCKELKRCVCERRAGMVKPESLHVPPETLGED

175/392

**FIGURE 167**

CGCAGGGGAGGACGCCCCGTTGCGCTAGCGCGTGCTCAGGAGTTGGTGTCTGCCTGCG  
CTCAGGATGAGGGGGAATCTGGCCCTGGTGGGCGTTCTAATCAGCCTGGCCTTCCTGTCA  
CTGCTGCCATCTGGACATCCTCAGCCGGCTGGCGATGACGCCTGCTCTGTGCAGATCCTC  
GTCCCTGGCCTCAAAGGGGATGCGGGAGAGAAGGGAGACAAAGGCGCCCCCGGACGGCCT  
GGAAGAGTCGGCCCCACGGGAGAAAAAGGAGACATGGGGGACAAAGGACAGAAAGGCAGT  
GTGGGTTCGTATGGAAAAATTGGTCCCATTGGCTCTAAAGGTGAGAAAGGAGATTCCGGT  
GACATAGGACCCCCCTGGTCCTAATGGAGAACCAGGCCTCCCATGTGAGTGCAGCCAGCTG  
CGCAAGGCCATCGGGGAGATGGACAACCAGGTCTCTCAGCTGACCAGCGAGCTCAAGTTC  
ATCAAGAATGCTGTGCGCGGTGTGCGCGAGACGGAGAGCAAGATCTACCTGCTGGTGAAG  
GAGGAGAAGCGCTACGCGGACGCCCAGCTGTCTGCCAGGGCCGCGGGGGCACGCTGAGC  
ATGCCCAAGGACGAGGCTGCCAATGGCCTGATGGCCGCATACCTGGCGCAAGCCGGCCTG  
GCCCCGTGTCTTCATCGGCATCAACGACCTGGAGAAGGAGGGCGCCTTCGTGTACTCTGAC  
CACTCCCCCATGCGGACCTTCAACAAGTGGCGCAGCGGTGAGCCCAACAATGCCTACGAC  
GAGGAGGACTGCGTGGAGATGGTGGCCTCGGGCGGCTGGAACGACGTGGCCTGCCACACC  
ACCATGTACTTCATGTGTGAGTTTGACAAGGAGAACATGTGAGCCTCAGGCTGGGGCTGC  
CCATTGGGGGCCCCACATGTCCCTGCAGGGTTGGCAGGGACAGAGCCCAGACCATGGTGC  
CAGCCAGGGAGCTGTCCCTCTGTGAAGGGTGGAGGCTCACTGAGTAGAGGGCTGTTGTCT  
AACTGAGAAAATGGCCTATGCTTAAGAGGAAAATGAAAGTGTTCTGGGGTGCTGTCTC  
TGAAGAAGCAGAGTTTCATTACCTGTATTGTAGCCCCAATGTCATTATGTAATTATTACC  
CAGAATTGCTCTTCCATAAAGCTTGTGCCTTTGTCCAAGCTATACAATAAAATCTTTAAG  
TAGTGCAGTAGTTAAGTCCAAAAAAAAAAAAAAAAAAAAA

176/392

**FIGURE 168**

MRGNLALVGVLI SLAFLSLLPSGHPQPAGDDACSVQILVPGLKGDAGEKGDKGAPGRPGR  
VGPTGEKGDMGDKGQKGSVGRHGKIGPIGSKGEKGDSDIGPPGPNGEPGLPCECSQLRK  
AIGEMDNQVSQLTSELKFIKNAVAGVRETESKIYLLVKEEKRYADAQLSCQGRGGTLSMP  
KDEAANGLMAAYLAQAGLARVFIGINDLEKEGAFVYSDHSPMRTFNKWRSGEPNNAYDEE  
DCVEMVASGGWNDVACHTTMYFMCEFDKENM



177/392

**FIGURE 169**

AGTGA CTGCAGCCTTCCTAGATCCCCTCCACTCGGTTTCTCTCTTTGCAGGAGCACCGGC  
AGCACCAGTGTGTGAGGGGAGCAGGCAGCGGTCC TAGCCAGTTCCTTGATCCTGCCAGAC  
CACCCAGCCCCCGGCACAGAGCTGCTCCACAGGCACCATGAGGATCATGCTGCTATTAC  
AGCCATCCTGGCCTTCAGCCTAGCTCAGAGCTTTGGGGCTGTCTGTAAAGGAGCCACAGGA  
GGAGGTGGTTCTTGGCGGGGGCCGAGCAAGAGGGATCCAGATCTCTACCAGCTGCTCCA  
GAGACTCTTCAAAGCCACTCATCTCTGGAGGGATTGCTCAAAGCCCTGAGCCAGGCTAG  
CACAGATCCTAAGGAATCAACATCTCCCGAGAAACGTGACATGCATGACTTCTTTGTGGG  
ACTTATGGGCAAGAGGAGCGTCCAGCCAGAGGGAAAGACAGGACCTTTCTTACCTTCAGT  
GAGGGTTCCTCGGCCCCCTTCATCCCAATCAGCTTGGATCCACAGGAAAGTCTTCCCTGGG  
AACAGAGGAGCAGAGACCTTTATAAGACTCTCCTACGGATGTGAATCAAGAGAACGTCCC  
CAGCTTTGGCATCCTCAAGTATCCCCGAGAGCAGAATAGGTACTCCACTTCCGGACTCC  
TGGACTGCATTAGGAAGACCTCTTTCCCTGTCCCAATCCCCAGGTGCGCACGCTCCTGTT  
ACCCTTTCTCTTCCCTGTTCTTGTAACATTCTTGTGCTTTGACTCCTTCTCCATCTTTTC  
TACCTGACCCTGGTGTGGAACTGCATAGTGAATATCCCCAACCCCAATGGGCATTGACT  
GTAGAATACCCTAGAGTTCCTGTAGTGTCTACATTAAAAATATAATGTCTCTCTCTATT  
CCTCAACAATAAAGGATTTTTCATATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAA

178/392

## FIGURE 170

MRIMLLFTAILAFSLAQSFQAVCKEPQEEVVPGGGRSKRDPDLYQLLQRLFKSHSSLEGL  
LKALSQASTDPKESTSPKRDMDHDFVGLMGKRSVQPEGKTGPFLPSVRVPRPLHPNQLG  
STGKSSLGTEEQRPL

**Important features:**

**Signal peptide:**

amino acids 1-18

**Tyrosine kinase phosphorylation site:**

amino acids 36-45

**N-myristoylation site:**

amino acids 33-39, 59-65

**Amidation site:**

amino acids 90-94

**Leucine zipper pattern:**

amino acids 43-65

**Tachykinin family signature:**

amino acids 86-92

179/392

**FIGURE 171**

TGGCCTCCCCAGCTTGCCAGGCACAAGGCTGAGCGGGAGGAAGCGAGAGGCATCTAAGCA  
GGCAGTGTTTTGCCTTCACCCCAAGTGACCATGAGAGGTGCCACGCGAGTCTCAATCATG  
CTCCTCCTAGTAACCTGTGTCTGACTGTGCTGTGATCACAGGGGCCTGTGAGCGGGATGTC  
CAGTGTGGGGCAGGCACCTGCTGTGCCATCAGCCTGTGGCTTCGAGGGCTGCGGATGTGC  
ACCCCGCTGGGGCGGGAAGGCGAGGAGTGCCACCCCGGCAGCCACAAGGTCCCTTCTTC  
AGGAAACGCAAGCACACACCTGTCTTGTCTTGCCCAACCTGCTGTGCTCCAGGTTCCTCG  
GACGGCAGGTACCGCTGCTCCATGGACTTGAAGAACATCAATTTTTTAGGCGCTTGCCCTGG  
TCTCAGGATACCCACCATCCTTTCTCTGAGCACAGCCTGGATTTTTATTTCTGCCATGAA  
ACCCAGCTCCCATGACTCTCCAGTCCCTACACTGACTACCTGATCTCTCTTGTCTAGT  
ACGCACATATGCACACAGGCAGACATACTCCCATCATGACATGGTCCCCAGGCTGGCCT  
GAGGATGTCACAGCTTGAGGCTGTGGTGTGAAAGGTGGCCAGCCTGGTTCTCTTCCCTGCT  
TCAGGCTGCCAGAGAGGTGGTAAATGGCAGAAAGGACATTCCCCCTCCCCCTCCCCAGGTG  
ACCTGCTCTCTTTCTTGGGCCCTGCCCCCTCTCCCCACATGTATCCCTCGGTCTGAATTAG  
ACATTCCTGGGCACAGGCTCTTGGGTGCATTGCTCAGAGTCCCAGGTCCTGGCCTGACCC  
TCAGGCCCTTCACGTGAGGTCTGTGAGGACCAATTTGTGGGTAGTTTCATCTTCCCTCGAT  
TGGTTAACTCCTTAGTTTCAGACCACAGACTCAAGATTGGCTCTTCCCAGAGGGCAGCAG  
ACAGTCACCCCAAGGCAGGTGTAGGGAGCCAGGGAGGCCAATCAGCCCCCTGAAGACTC  
TGGTCCCAGTCAGCCTGTGGCTTGTGGCCTGTGACCTGTGACCTTCTGCCAGAATTGTCA  
TGCCTCTGAGGCCCCCTCTTACCACACTTTACCAGTTAACCCTGAAGCCCCCAATTCCC  
ACAGCTTTTCCATTAAAATGCAAATGGTGGTGGTTCAATCTAATCTGATATTGACATATT  
AGAAGGCAATTAGGGTGTTCCTTAAACAACCTCCTTTCCAAGGATCAGCCCTGAGAGCAG  
GTTGGTGACTTTGAGGAGGGCAGTCCTCTGTCCAGATTGGGGTGGGAGCAAGGGACAGGG  
AGCAGGGCAGGGGCTGAAAGGGGCACTGATTGAGACCAGGGAGGCAACTACACACCAACA  
TGCTGGCTTTAGAATAAAAGCACCAACTGAAAAAA

180/392

## FIGURE 172

MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCAI~~SLWLRGLRMCTPLGREGE~~EC  
HPGSHKVPFFRKRKHHTCPCLPNLLCSRFPDGRYRCSMDLKNINF

Signal peptide:  
amino acids 1-19

Tyrosine kinase phosphorylation site:  
amino acids 88-95

N-myristoylation sites:  
amino acids 33-39, 35-41, 46-52

181/392

**FIGURE 173**

AGCGCCCGGGCGTTCGGGGCGGTAAAAGGCCGGCAGAAAGGGAGGCACTTGAGAAATGCTCTT  
TCCTCCAGGACCCAAGTTTCTTCACCATGGGGATGTGGTCCATTGGTGCAGGAGCCCTGG  
GGGCTGCTGCCTTGGCATTGCTGCTTGCCAACACAGACGTGTTTCTGTCCAAGCCCCAGA  
AAGCGGCCCTGGAGTACCTGGAGGATATAGACCTGAAAACACTGGAGAAGGAACCAAGGA  
CTTTCAAAGCAAAGGAGCTATGGGAAAAAATGGAGCTGTGATTATGGCCGTGCGGAGGC  
CAGGCTGTTTCCTCTGTCTGAGAGGAAGCTGCGGATCTGTCTCCCTGAAAAGCATGTTGG  
ACCAGCTGGGCGTCCCCCTCTATGCAGTGGTAAAGGAGCACATCAGGACTGAAGTGAAGG  
ATTTCCAGCCTTATTTCAAAGGAGAAATCTTCCTGGATGAAAAGAAAAAGTTCTATGGTC  
CACAAAGGCGGAAGATGATGTTTATGGGATTTATCCGTCTGGGAGTGTGGTACAACCTTCT  
TCCGAGCCTGGAACGGAGGCTTCTCTGGAACCTGGAAGGAGAAGGCTTCATCCTTGGGG  
GAGTTTTCTGTGGTGGGATCAGGAAAGCAGGGCATTCTTCTTGAGCACCGAGAAAAAGAAT  
TTGGAGACAAAGTAAACCTACTTTCTGTTCTGGAAGCTGCTAAGATGATCAAACCACAGA  
CTTTGGCCTCAGAGAAAAAATGATTGTGTGAAACTGCCCAGCTCAGGGATAACCAGGGAC  
ATTCACCTGTGTTTCATGGGATGTATTGTTTCCACTCGTGTCCCTAAGGAGTGAGAAACCC  
ATTTATACTCTACTCTCAGTATGGATTATTAATGTATTTTAATATTCTGTTTAGGCCCAC  
TAAGGCAAAATAGCCCCAAAACAAGACTGACAAAAATCTGAAAAACTAATGAGGATTATT  
AAGCTAAAACCTGGGAAATAGGAGGCTTAAAAATTGACTGCCAGGCTGGGTGCAGTGGCTC  
ACACCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGAGCAAGTCACTTGAGGTCGGGAGT  
TCGAGACCAGCCTGAGCAACATGGCGAAACCCCGTCTCTACTAAAAATACAAAAATCACC  
CGGGTGTGGTGGCAGGCACCTGTAGTCCCAGCTACCCGGGAGGCTGAGGCAGGAGAATCA  
CTTGAACCTGGGAGGTGGAGGTTGCGGTGAGCTGAGATCACACCCTGTATTCCAGCCTG  
GGTGA CTGAGACTCTAACTAA

182/392

**FIGURE 174**

MSFLQDPSFFTMGMWSIGAGALGAAALALLANTDVFLSKPQKALEYLEDIDLKTLEKE  
PRTFKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTE  
VKDFQPYFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGWYNFFRAWNGGFSGNLEGEGLI  
LGGVFVVGSGKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK

183/392

## FIGURE 175

GACAGTGGAGGGCAGTGGAGAGGACCGCGCTGTCCTGCTGTCACCAAGAGCTGGAGACAC  
CATCTCCCACCGAGAGTCATGGCCCCATTGGCCCTGCACCTCCTCGTCCTCGTCCCCATC  
CTCCTCAGCCTGGTGGCCTCCCAGGACTGGAAGGCTGAACGCAGCCAAGACCCCTTCGAG  
AAATGCATGCAGGATCCTGACTATGAGCAGCTGCTCAAGGTGGTGACCTGGGGGCTCAAT  
CGGACCCTGAAGCCCCAGAGGGTGATTGTGGTTGGCGCTGGTGTGGCCGGGCTGGTGGCC  
GCCAAGGTGCTCAGCGATGCTGGACACAAGGTCACCATCCTGGAGGCAGATAACAGGATC  
GGGGGCCGCATCTTCACCTACCGGGACCAGAACACGGGCTGGATTGGGGAGCTGGGAGCC  
ATGCGCATGCCCAGCTCTCACAGGATCCTCCACAAGCTCTGCCAGGGCCTGGGGCTCAAC  
CTGACCAAGTTACCCAGTACGACAAGAACAGTGGACGGAGGTGCACGAAGTGAAGCTG  
CGCAACTATGTGGTGGAGAAGGTGCCCCGAGAAGCTGGGCTACGCCTTGCGTCCCCAGGAA  
AAGGGCCACTCGCCCCGAAGACATCTACCAGATGGCTCTCAACCAGGCCCTCAAAGACCTC  
AAGGCACTGGGCTGCAGAAAGGCGATGAAGAAGTTTGAAGGCACACGCTCTTGGAATAT  
CTTCTCGGGGAGGGGAACCTGAGCCGGCCGGCGTGACGCTTCTGGGAGACGTGATGTCC  
GAGGATGGCTTCTTCTATCTCAGCTTCGCCGAGGCCCTCCGGGCCACAGCTGCCTCAGC  
GACAGACTCCAGTACAGCCGCATCGTGGGTGGCTGGGACCTGCTGCCGCGCGCTGCTG  
AGCTCGCTGTCCGGGCTTGTGCTGTTGAACGCGCCCCGTGGTGGCGATGACCCAGGGACCG  
CACGATGTGCACGTGCAGATCGAGACCTCTCCCCGGCGCGGAATCTGAAGGTGCTGAAG  
GCCGACGTGGTGCTGCTGACGGCGAGCGGACCGGCGGTGAAGCGCATCACCTTCTCGCCG  
CCGCTGCCCCGCCACATGCAGGAGGCGCTGCGGAGGCTGCACTACGTGCCGGCCACCAAG  
GTGTTCTTAAGCTTCCGCAGGCCCTTCTGGCGCGAGGAGCACATTGAAGGCGGCCACTCA  
AACACCGATCGCCCGTCGCGCATGATTTTCTACCCGCCCGCGCGAGGGCGCGCTGCTG  
CTGGCCTCGTACACGTGGTCCGACGCGCGGCGAGCGTTCGCCGGCTTGAGCCGGGAAGAG  
GCGTTGCGCTTGGCGCTCGACGACGTGGCGGCATTGCACGGGCCTGTCGTGCGCCAGCTC  
TGGGACGGCACCCGGCGTCGTCAAGCGTTGGGCGGAGGACCAGCACAGCCAGGGTGGCTTT  
GTGGTACAGCCGCCGGCGCTCTGGCAAACCGAAAAGGATGACTGGACCGTCCCTTATGGC  
CGCATCTACTTTGCCGGCGAGCACACCGCCTACCCGCACGGCTGGGTGGAGACGGCGGTC  
AAGTCGGCGCTGCGCGCCGCCATCAAGATCAACAGCCGGAAGGGGCTGCATCGGACACG  
GCCAGCCCCGAGGGGCACGCATCTGACATGGAGGGGCAGGGGCATGTGCATGGGGTGGCC  
AGCAGCCCCCTCGCATGACCTGGCAAAGGAAGAAGGCAGCCACCCTCCAGTCCAAGGCCAG  
TTATCTCTCAAAACACGACCCACACGAGGACCTCGCATTAAAGTATTTTCGGAAAAAAA  
AAA

184/392

**FIGURE 176**

MAPLALHLLVLPILLSLVSQDWKAERSQDPFEKCMQDPDYEQLLKVVVTWGLNRTLKPQ  
RVIVVGAGVAGLVAAKVLSDAGHKVTILEADNRIGGRIFTYRDQNTGWIGELGAMRMPSS  
HRILHKLCQGLGLNLTKFTQYDKNTWTEVHEVKLRNYVVEKVPEKLGALRPQEKGHSPE  
DIYQMALNQALKDLKALGCRKAMKKFERHTLLEYLLGEGNLSRPAVQLLGDVMSDGGFFY  
LSFAEALRAHSCLSDRLQYSRIVGGWDLPRALLSSLSGLVLLNAPVVAMTQGPHDVHVQ  
IETSPPARNLKVLKADVLLTASGPAVKRITFSPPLPRHMQEALRRLHYVPATKVFLSFR  
RPFWREEHIEGGHSNTDRPSRMIFYPPPREGALLASYTWSDAFAAGLSREEALRLAL  
DDVAALHGPVVRQLWDGTGVVKRWAEDQHSQGGFVVQPPALWQTEKDDWTVPYGRIYFAG  
EHTAYPHGWVETAVKSALRAAIKINSRKGPASDTASPEGHASDMEGQGHVHGVASSPSHD  
LAKEEGSHPPVQGQLSLQNTTHTRTSH

**Signal peptide:**  
amino acids 1-21



185/392

**FIGURE 177**

CCGGGGAGGGGAGGGCCCCGTCCCGCCCCCTCCCCGTCTCTCCCCGCCCCCTCCCCGTCCCTC  
CCGCCGAAGCTCCGTCCCGCCCCGCGGGCCGGCTCCGCCCTCACCTCCCGGCCGCGGCTGC  
CCTCTGCCCGGGTTGTCCAAGATGGAGGGCGCTCCACCGGGGTCGCTCGCCCTCCGGCTC  
CTGCTGTTTCGTGGCGCTACCCGCCTCCGGCTGGCTGACGACGGGCGCCCCGAGCCGCCG  
CCGCTGTCCGGAGCCCCACAGGACGGCATCAGAATTAATGTAAC TACACTGAAAGATGAT  
GGGGACATATCTAAACAGCAGGTTGTTCTTAACATAACCTATGAGAGTGGACAGGTGTAT  
GTAAATGACTTACCTGTAAATAGTGGTGTAACCCGAATAAGCTGTCAGACTTTGATAGTG  
AAGAATGAAAATCTTGAAAATTTGGAGGAAAAAGAATATTTTGGAATTGTCAGTGTAAGG  
ATTTTAGTTCATGAGTGGCCTATGACATCTGGTTCCAGTTTGCAACTAATTGTCATTCAA  
GAAGAGGTAGTAGAGATTGATGGAAAACAAGTTCAGCAAAAGGATGTCACTGAAATTGAT  
ATTTTAGTTAAGAACCGGGGAGTACTCAGACATTCAAAC TATACCCTCCCTTTGGAAGAA  
AGCATGCTCTACTCTATTTCTCGAGACAGTGACATTTTATTTACCCTTCCTAACCTCTCC  
AAAAAAGAAAGTGTTAGTTCAGTGCAAACCACTAGCCAGTATCTTATCAGGAATGTGGAA  
ACCACTGTAGATGAAGATGTTTTACCTGGCAAGTTACCTGAAACTCCTCTCAGAGCAGAG  
CCGCCATCTTCATATAAGGTAATGTGTCACTGGATGGAAAAGTTTAGAAAAGATCTGTGT  
AGGTTCTGGAGCAACGTTTTCCAGTATTCTTTCAGTTTTTGAACATCATGGTGGTTGGA  
ATTACAGGAGCAGCTGTGGTAATAACCATCTTAAAGGTGTTTTTCCAGTTTCTGAATAC  
AAAGGAATTCTTCAGTTGGATAAAAGTGGACGTCATACCTGTGACAGCTATCAACTTATAT  
CCAGATGGTCCAGAGAAAAGAGCTGAAAACCTTGAAGATAAAACATGTATTTAAACGCC  
ATCTCATATCATGGACTCCGAAGTAGCCTGTTGCCTCCAAATTTGCCACTTGAATATAAT  
TTTCTTTAAATCGTT

186/392

**FIGURE 178**

MEGAPPGSLALRLLLFVALPASGWLTTGAPEPPPLSGAPQDGIRINVTTLKDDGDISKQQ  
VVLNITYESGQVYVNDLPVNSGVTRISCQTLIVKNENLENLEEKEYFGIVSVRILVHEWP  
MTSGSSLQLIVIQEEVVEIDGKQVQQKDVTEIDILVKNRGVLRHSNYTLPLEESMLYSIS  
RDS DILFTLPNL SKKESVSSLQTTSQYLIRNVETTVD EDVLP GKLPETPLRAEPPSSYKV  
MCQWMEKFRKDL CRFWSNVFPVFFQFLNIMVVGITGA AVVITILKVFFPVSEYKGILQLD  
KVDVIPVTAINLYPDGPEKRAENLEDKTCI

Signal peptide:

1-23

Transmembrane domain:

266-284

Leucine zipper pattern:

155-177

N-glycosylation site:

46-50, 64-68, 166-170, 191-195

Motif name: N-myristoylation site:

3-9, 42-48, 273-279

187/392

**FIGURE 179**

CTCCTTAGGTGGAAACCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATAC  
GTCCCCGGGCAGGGGTGACAACAGGTGTCTATCTTTTGATCTCGTGTGTGGCTGCCTTCC  
TATTTCAAGGAAAGACGCCAAGGTAATTTTGACCCAGAGGAGCAATGATGTAGCCACCTC  
CTAACCTTCCCTTCTTGAACCCCCAGTTATGCCAGGATTTACTAGAGAGTGTCAACTCAA  
CCAGCAAGCGGCTCCTTCGGCTTAACCTTGTGGTTGGAGGAGAGAACCCTTGTGGGGCTGC  
GTTCTCTTAGCAGTGCTCAGAAAGTGACTTGCCCTGAGGGTGGACCAGAAGAAAGGAAAGGT  
CCCCTCTTGCTGTTGGCTGCACATCAGGAAGGCTGTGATGGGAATGAAGGTGAAAACCTTG  
GAGATTTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCCTTTAAAAGTAGAGAAGCT  
GCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAG  
CAGCTCCGGGGGGCCCCAAACGCATGCTTCTGTGGTCTAGCCCAGGGAAGCCCTTCCGTG  
GGGGCCCCGGCTTTGAGGGATGCCACCGGTTCTGGACGCATGGCTGATTCTGAATGATG  
ATGGTTCGCCGGGGGGCTGCTTGCCTGGATTTCCCGGGTGGTGGTTTGTGGTGTCTCCTC  
TGCTGTGCTATCTGTCTGTACATGTTGGCCTGCACCCCCAAAAGGTGACGAGGAGCAG  
CTGGCACTGCCCCAGGGCCAACAGCCCCACGGGGAAGGAGGGGTACCAGGCCGTCTTCAG  
GAGTGGGAGGAGCAGCACCGCAACTACGTGAGCAGCCTGAAGCGGCAGATCGCACAGCTC  
AAGGAGGAGCTGCAGGAGAGGAGTGAGCAGCTCAGGAATGGGCAGTACCAAGCCAGCGAT  
GCTGCTGGCCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAGGCCGACCTCCTGGCC  
TTCCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGGCCACAGAG  
TATGCAGCAGTGCCTTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGACTGGC  
CTTACCCGCCACCCCGAGGAGAAGCCTGTGAGGAAGGACAAGCGGGATGAGTTGGTGGAA  
GCCATTGAATCAGCCTTGGAGACCTGAACAATCCTGCAGAGAACAGCCCCAATCACCGT  
CCTTACACGGCCTCTGATTTCATAGAAGGGATCTACCGAACAGAAAGGGACAAAGGGACA  
TTGTATGAGCTCACCTTCAAAGGGGACCACAAACACGAATTCAAACGGCTCATCTTATTT  
CGACCATTGAGCCCCATCATGAAAGTGAAAATGAAAAGCTCAACATGGCCAACACGCTT  
ATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGGACAAGTTCGGGCAGTTCATGCAGAAT  
TTCAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGG  
AAAGAAGAAATAAATGAAGTCAAAGGAATACTTGAAAACACTTCCAAAGCTGCCAACTTC  
AGGAACTTTACCTTCATCCAGCTGAATGGAGAATTTTCTCGGGGAAAGGGACTTGATGTT  
GGAGCCCGCTTCTGGAAGGGAAGCAACGTCCTTCTCTTTTCTGTGATGTGGACATCTAC  
TTCACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAGGTATTT  
TATCCAGTTCTTTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTC  
CCTCCCTTGGAACAGCAGCTGGTCATAAAGAAGGAACTGGATTTTGGAGAGACTTTGGA  
TTTGGGATGACGTGTCAGTATCGGTGAGACTTCATCAATATAGGTGGGTTTGATCTGGAC  
ATCAAAGGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTATCTCCACAGCAACCTC  
ATAGTGGTACGGACGCCTGTGCGAGGACTTCCACCTCTGGCATGAGAAGCGCTGCATG  
GACGAGCTGACCCCCGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCA  
TCCCACGGCCAGCTGGGCATGCTGGTGTTCAGGCACGAGATAGAGGCTCACCTTCGCAA  
CAGAAACAGAAGACAAGTAGCAAAAAACATGAACTCCAGAGAAGGATTGTGGGAGACA  
CTTTTTCTTTCTTTTGCAATTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAA  
GGACGACAAAAGAATTGGACTGATGGGTGAGAGATGAGAAAAGCCTCCGATTTCTCTGT  
TGGGCTTTTTTACAACAGAAATCAAATCTCCGCTTTGCCTGCAAAAGTAACCCAGTTGCA  
CCCTGTGAAGTGTCTGACAAAGGCAGAATGCTTGTGAGATTATAAGCCTAATGGTGTGGA  
GGTTTTGATGGTGTTTACAATACACTGAGACCTGTTGTTTTGTGTGCTCATTGAAATATT  
CATGATTTAAGAGCAGTTTGTAAAAAATTCATTAGCATGAAAGGCAAGCATATTTCTCC  
TCATATGAATGAGCCTATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAAATATCAGAAGG  
CAGGAGAGGAGATAGGCTTATTATGATACTAGTGAGTACATTAAGTAAAATAAATGGAC  
CAGAAAAGAAAAGAAACCATAAATATCGTGTCTATTTTTCCCAAGATTAACCAAAAATA

188/392

ATCTGCTTATCTTTTTGGTTGTCCTTTTAACTGTCTCCGTTTTTTTCTTTTATTTAAAAA  
TGCACTTTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCACTTTGCAAGCCTT  
ACAAGAGAGCACAAAGTTGGCCTACATTTTTATATTTTTTAAGAAGATACTTTGAGATGCA  
TTATGAGAACTTTTCAAGCATCAAATTGATGCCATATCCAAGGACATGCCAAATG  
CTGATTCTGTGAGGCACTGAATGTCAGGCATTGAGACATAGGGAAGGAATGGTTTGTACT  
AATACAGACGTACAGATACTTTCTCTGAAGAGTATTTTCGAAGAGGAGCAACTGAACACT  
GGAGGAAAAGAAAATGACACTTTCTGCTTTACAGAAAAGGAAACTCATTGAGACTGGTGA  
TATCGTGATGTACCTAAAAGTCAGAAACCACATTTTCTCCTCAGAAGTAGGGACCGCTTT  
CTTACCTGTTTAAATAAACCAAAGTATACCGTGTGAACCAAACAATCTCTTTTCAAAACA  
GGGTGCTCCTCCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAAATATATATATATA  
TATATATATTGTGAAAGATCAATCCATCTGCCAGAATCTAGTGGGATGGAAGTTTTTGCT  
ACATGTTATCCACCCAGGCCAGGTGGAAGTAACTGAATTATTTTTTAAATTAAGCAGTT  
CTACTCAATCACCAAGATGCTTCTGAAAATTGCATTTTATTACCATTTCAACTATTTTT  
TAAAAATAAATACAGTTAACATAGAGTGGTTTCTTCATTCATGTGAAAATTATTAGCCAG  
CACCAGATGCATGAGCTAATTATCTCTTTGAGTCCTTGCTTCTGTTTGCTCACAGTAAAC  
TCATTGTTTAAAAGCTTCAAGAACATTCAAGCTGTTGGTGTGTTAAAAAATGCATTGTAT  
TGATTTGTACTGGTAGTTTATGAAATTTAATTAAAAACACAGGCCATGAATGGAAGGTGGT  
ATTGCACAGCTAATAAAATATGATTTGTGGATATGAA

189/392

**FIGURE 180**

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPGKEGYQAV  
LQEWEEQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGGLDRSPPEKTQADL  
LAFLHSQVDKAEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPPEKPVKDKRDEL  
VEAIESALETLNPAENSPNHRPYTASDFIEGIYRTERDKGTLTYELTFKGDHKHEFKRLI  
LFRPFSPIMKVKNEKLNMANLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVY  
FGKEEINEVKGILENTSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVD  
IYFTSEFLNTCRLNTQPGKKVFYPVLFSQYNPGIIYGHDAVPPLEQQLVIKKETGFWRD  
FGFGMTCQYRSDFINIGGFDLDIKGWGGEDVHLYRKYLHSNLIVVRTPVRLFHLWHEKR  
CMDELTPEQYKCMQSKAMNEASHGQLGMLVFRHEIEAHLRKQKQKTSSKKT

190/392

**FIGURE 181**

CGTCTCTGCGTTCCGCCATGCGTCCCGGGGCGCCAGGGCCACTCTGGCCTCTGCCCTGGGG  
GGCCCTGGCTTGGGCGGTGGGCTTCGTGAGCTCCATGGGCTCGGGGAACCCCGCGCCCGG  
TGGTGTGTTGCTGGCTCCAGCAGGGCCAGGAGGCCACCTGCAGCCTGGTGCTCCAGACTGA  
TGTCACCCGGGCGAGTGCTGTGCCTCCGGCAACATTGACACCGCCTGGTCCAACCTCAC  
CCACCCGGGGAACAAGATCAACCTCCTCGGCTTCTTGGGCCTTGTCCTACTGCCTTCCCTG  
CAAAGATTGCTGCGACGGCGTGAGTGCGGCCCCGGGCAAGGCGTGCCGCATGCTGGGGGG  
CCGCCCCGCGCTGCGAGTGCGCGCCCCGACTGCTCGGGGCTCCCGGCGCGGCTGCAGGTCTG  
CGGCTCAGACGGCGCCACCTACCGCGACGAGTGCGGAGCTGCGCGCCGCGCGCTGCCGCGG  
CCACCCGGACCTGAGCGTCATGTACCGGGGCGGCTGCCGCAAGTCCTGTGAGCACGTGGT  
GTGCCCGCGGGCCACAGTCGTGCGTCGTGGACCAGACGGGCAGCGCCCACTGCGTGCGTGTG  
TCGAGCGGCGCCCTGCCCTGTGCCCTCCAGCCCCGGCCAGGAGCTTTCGCGCAACAACAA  
CGTCACCTACATCTCCTCGTGCCACATGCGCCAGGCCACCTGCTTCTGGGCGCTCCAT  
CGGCGTGCGCCACGCGGGCAGCTGCGCAGGCACCCCTGAGGAGCCGCCAGGTGGTGAGTC  
TGCAGAAGAGGAAGAGAACTTCGTGTGAGCCTGCAGGACAGGCCTGGGCCTGGTGCCCGA  
GGCCCCCATCATCCCTGTTATTTATTGCCACAGCAGAGTCTAATTTATATGCCACGGA  
CACTCCTTAGAGCCCGGATTTCGGACCCTTGGGGATCCCAGAACCTCCCTGACGATATCC  
TGGAAGGACTGAGGAAGGGAGGCCTGGGGGCGGCTGGTGGGTGGGATAGACCTGCGTTC  
CGGACACTGAGCGCCTGATTTAGGGCCCTTCTCTAGGATGCCCCAGCCCTACCTAAGA  
CCTATTGCCGGGGAGGATTCCACACTTCCGCTCCTTTGGGGATAAACCTATTAATTATTG  
CTACTATCAAGAGGGGCTGGGCATTCTCTGCTGGTAATTCCTGAAGAGGCATGACTGCTTT  
TCTCAGCCCCAAGCCTCTAGTCTGGGTGTGTACGGAGGGTCTAGCCTGGGTGTGTACGGA  
GGGTCTAGCCTGGGTGAGTACGGAGGGTCTAGCCTGGGTGAGTACGGAGGGTCTAGCCTG  
GGTGAGTACGGAGGGTCTAGCCTGGGTGTGTATGGAGGATCTAGCCTGGGTGAGTATGGA  
GGGTCTAGCCTGGGTGAGTATGGAGGGTCTAGCCTGGGTGTGTATGGAGGGTCTAGCCTG  
GGTGAGTATGGAGGGTCTAGCCTGGGTGTGTATGGAGGGTCTAGCCTGGGTGAGTATGGA  
GGGTCTAGCCTGGGTGTGTACGGAGGGTCTAGTCTGAGTGCGTGTGGGGACCTCAGAACA  
CTGTGACCTTAGCCCAGCAAGCCAGGCCCTTCATGAAGGCCAAGAAGGCTGCCACCATT  
CCTGCCAGCCCAAGAACTCCAGCTTCCCCACTGCCTCTGTGTGCCCCCTTTCGCTCCTGTG  
AAGGCCATTGAGAAATGCCCAGTGTGCCCCCTGGGAAAGGGCACGGCCTGTGCTCCTGAC  
ACGGGCTGTGCTTGGCCACAGAACCACCCAGCGTCTCCCCTGCTGCTGTCCACGTAGTT  
CATGAGGCAACGTGCGCTGGTCTCAGACGTGGAGCAGCCAGCGGCAGCTCAGAGCAGGGC  
ACTGTGTCCGGCGGAGCCAAGTCCACTCTGGGGGAGCTCTGGCGGGGACCACGGGCCACT  
GCTCACCCACTGGCCCCGAGGGGGGTGTAGACGCCAAGACTCACGCATGTGTGACATCCG  
GAGTCCTGGAGCCGGGTGTCCCAGTGGCACCCTAGGTGCCTGCTGCCTCCACAGTGGGG  
TTCACACCCAGGGCTCCTTGGTCCCCCACAACCTGCCCCGGCCAGGCCTGCAGACCCAGA  
CTCCAGCCAGACCTGCCTCACCCACCAATGCAGCCGGGGCTGGCGACACCAGCCAGGTGC  
TGGTCTTGGGCCAGTTCTCCACGACGGCTCACCTCCCCCTCCATCTGCGTTGATGCTCA  
GAATCGCTACCTGTGCCTGCGTGTAACCACAGCCTCAGACCAGCTATGGGGAGAGGAC  
AACACGGAGGATATCCAGCTTCCCCGGTCTGGGGTGAGGAATGTGGGGAGCTTGGGCATC  
CTCCTCCAGCCTCCTCCAGCCCCCAGGCAGTGCCTTACCTGTGGTGCCAGAAAAGTGCC  
CCTAGGTTGGTGGGTCTACAGGAGCCTCAGCCAGGCAGCCACCCACCCCTGGGGCCCTG  
CCTCACCAAGGAAATAAAGACTCAAGCCATAAAAAAA

191/392

## FIGURE 182

MRPGAPGPLWPLPWGALAWAVGVSSMGSGNPAPGGVCWLQQGQEATCSLVLQTDVTRAE  
CCASGNIDTAWSNLTHPGNKINLLGFLGLVHCLPCKDSCDGVECGPGKACRMLGGRPRCE  
CAPDCSGLPARLQVCGSDGATYRDECELRAARCRGHPDLSVMYRGRCRKSCHEVVCPRPQ  
SCVVDQTGSAHCVVCRAAPCFVPSSPGQELCGNNNVITYISSCHMRQATCFLGRSIGVRHA  
GSCAGTPEEPPGGESAEEEENFV

**Important features:**

**Signal peptide:**

amino acids 1-20

**N-glycosylation sites:**

amino acids 73-77, 215-219

**Osteonectin domain proteins:**

amino acids 97-130, 169-202

192/392

**FIGURE 183**

CACTCATTCAATTCCAAAGGGTCTCTCAAGGCAATGGTAATGTGCAAGGAGGTGATACCTA  
AATGAATGACCAAAAGAACATGCTTCTGCTTTTGTGTGTCTCCTACATTTTAGACATTTG  
TTTGTTCCTCTTGGTAGCCTTTAAATTCCTTGAAGCCCAGGACCATGTCTCACTTACCTT  
TGTGTTTCCACTAACTAGTCTACCTCCTGGAATTGGCAGATACTCAGTGAAAAGCCTGTGA  
AATAAGTGATGTCTATTTCTAGCATATTATTCTGAGATTTAATGATAGATTTAGTGATTG  
AATGAGATTTCCATTTTCAAATACAGCAAAGCATAACTATTTTCATTCATTCAATTCA  
TTCAACTTCATTCTCAAATTAGGTCTGAGTTAACTAATAATTACCTTTGAAATGTGTG  
GGTTATTTGAGGCAATCAGGTGGTGACATTGAGCTCTCAGCCAGAGTTTGTTCCTGGAAT  
TGATTCAGTTCCATTGCATTGATTTTTGTTCTCAGAAGCCAAGGTTTCCCATGAAAAATC  
ATTCCCACCTTGAATTGGGCTGTGATTCTTGCTGCGTTTAAAGTAAAGGAAGCCTCTTGGTT  
CTAGTCTCTGCAAACCTTACACACTGAACTGGGACAAGTTTTTGTTCAGTAATGGCTGGG  
AAAAGAGGAACCTTTCATTTTATTTCAGAAGTCAAAAACAAAGGCCTCCAGCCACCTGGGA  
GATGTTTTGTTGTCAGACACCAGCCTGGCTCTGTCTTTATGCCTAACAATTGAGCATCCAG  
TCTTCTTTGTGCTGGGACCATTGCTCAGCTCTGCAAGGGGAAAAGAGGGAGAAAAGCCAGA  
GCTGCCAGGCTTCTTGCACTGGGGCCGGGGAGGGTTCTGGGAAGCAGGTGCTCTCTGG  
CTTCTTGGTACGTGAGGCTCTCGGAGCTGCCTCTCCTCTGACCCTCAGGTCCTCACCAG  
TTTGCTCCAGGAGTATATTGAAAACATACCCAGTGCTCTCTCAAGCACCCACTGCTTAGA  
GGGCCCAGATTTCTTTTCTTCTTTCCCTTGCAAGAGCTGGAGACTGCATCGGGCATCTGG  
TGTTTAAACTAAACAGGAAAAGTGAATAAGGTCCACAGTGCTCATTGTGTAGACTAGCT  
GCCCTCCGATGGGTGCTCTGATTATCAGTGCTTCCAGTGCCAGGGCCTGTCACTAAACAGG  
CCTCACTTCCTCCTTGGGGGCTTTCCCATGGGAGGTGTGGCTTTTTACTCTACATGGAAA  
TGACTCTCTGCAGCCACAGAACACAGTCATTTTCTGAATTATCCAGTCTCTCATGCGCC  
CTGGATTCTCCAGATGCCTTATATCTCTTGTGCAAAGTTGTCTAAAATTTGGTTCCAG  
CTTCCAAGCCTTGCCTTTTGGCCTTCCCTGGAAGTATTTTTGTGATGAGTCGTCTGTCA  
TATTCTCTAAAATGATTTGCTTTTTGTTCCTTTTCATTCCTATTTCCACCCACATATACA  
CACATGCTTCTTAACTTAGGGGATTACATGCCAATAAATCTATTGTTGAAAATGCACTAA  
TACTATCGCAAAGACGAAAATTACAGGCTGAACCGTTGTAAGTCCATATGCTCCTCAAC  
TTACATGTGTGATGGAGTTATGCCCAAATAAGTCCATCGTCAAGTTGAAAATCAAAATC  
AAGCCATCTTAGGTTGAGGACCATTTGTTTGTACCTCCAAAGATGTCATATCTTTAAACA  
TACTCCCTAGCTTTTCTTTTACTTTTATTTTGAAGTAATTATAGAATCACAGAAAGTT  
GCAAAAAA



193/392

## FIGURE 184

MGALIISGSSAGPVTQASLPPWGLSHGRCGFLLYMENTLCSHRTQSFSELSQSLMRPGF  
LQMPYISCAKLSKIWFPAKPCLLAFLEVFLLMSRLSLFSKMICFLFLSFLFPPHIYTHAS

Important features of the protein:

Signal peptide:

amino acids 1-41

Transmembrane domain:

amino acids 88-107

Casein kinase II phosphorylation site:

amino acids 47-50

N-myristoylation site:

amino acids 24-29

194/392

**FIGURE 185**

AACTCAAACCTCCTCTCTCTGGGAAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCA  
GGGTGTTGGAGCCCTCGGTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTC  
ATGTATGGCAAGAGCTCTACTCGTGCGGTGCTTCTTCTCCTTGGCATAACAGCTCACAGCT  
CTTTGGCCTATAGCAGCTGTGGAAATTTATACCTCCCGGGTGCTGGAGGCTGTTAATGGG  
ACAGATGCTCGGTAAAAATGCACTTTCTCCAGCTTTGCCCCCTGTGGGTGATGCTCTAACA  
GTGACCTGGAATTTTCGTCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCAC  
ATAGATCCCTTCCAACCCATGAGTGGGCGGTTTAAGGACCGGGTGCTTGGGATGGGAAT  
CCTGAGCGGTACGATGCCTCCATCCTTCTCTGGAAACTGCAGTTCGACGACAATGGGACA  
TACACCTGCCAGGTGAAGAACCCACCTGATGTTGATGGGGTGATAGGGGAGATCCGGCTC  
AGCGTCGTGCACACTGTACGCTTCTCTGAGATCCACTTCCTGGCTCTGGCCATTGGCTCT  
GCCTGTGCACTGATGATCATAATAGTAATTGTAGTGGTCCTCTTCCAGCATTACCGGAAA  
AAGCGATGGGCCGAAAGAGCTCATAAAGTGGTGAGATAAAATCAAAAGAAGAGGAAAGG  
CTCAACCAAGAGAAAAAGGTCTCTGTTTATTTAGAAAGACACAGACTTACAATTTTAGATG  
GAAGCTGAGATGATTTCCAAGAACAAAGAACCCTAGTATTTCTTGAAGTTAATGGAACTT  
TTCTTTGGCTTTTCCAGTTGTGACCCGTTTTCCAACCAGTTCTGCAGCATATTAGATTCT  
AGACAAGCAACACCCCTCTGGAGCCAGCACAGTGCTCCTCCATATCACCAGTCATACACA  
GCCTCATTATTAAGGTCTTATTTAATTTCAAGAGTGTAATTTTTTCAAGTGCTCATTAGG  
TTTTATAAACAAGAAGCTACATTTTGGCCCTTAAGACACTACTTACAGTGTTATGACTTG  
TATACACATATATTGGTATCAAAGGGGATAAAAGCCAATTTGTCTGTTACATTTCTTTC  
ACGTATTTCTTTTAGCAGCACTTCTGCTACTAAAGTTAATGTGTTTACTCTCTTTCCTTC  
CCACATTCTCAATTAAAAGGTGAGCTAAGCCTCCTCGGTGTTTCTGATTAACAGTAAATC  
CTAAATTCAAACGTAAATGACATTTTATTTTATGTCTCTCCTTAACATATGAGACAC  
ATCTTGTTTTACTGAATTTCTTTCAATATTCAGGTGATAGATTTTTTGTCTG

195/392

## FIGURE 186

MYGKSSTRAVLLLLGIQLTALWPAAVEIYTSRVLEAVNGTDARLKCTFSSSFAPVGDALT  
VTWNFRPLDGGPEQFVFYYHIDPFQPMSEGRFKDRVSWDGNPERYDASILLWKLQFDDNGT  
YTCQVKNPPDVDGVIGEIRLSVVHTVRFSEIHFLALAIGSACALMIIIVIVVVVLFQHYRK  
KRWAEARAHKVVEIKSKEEERLNQEKVSVYLETD

196/392

**FIGURE 187**

GCATTTTTGTCTGTGCTCCCTGATCTTCAGGTCACCACCATGAAGTTCTTAGCAGTCCTG  
GTACTCTTGGGAGTTTCCATCTTTCTGGTCTCTGCCCAGAATCCGACAACAGCTGCTCCA  
GCTGACACGTATCCAGCTACTGGTCCTGCTGATGATGAAGCCCCTGATGCTGAAACCACT  
GCTGCTGCAACCACTGCGACCACTGCTGCTCCTACCACTGCAACCACCGCTGCTTCTACC  
ACTGCTCGTAAAGACATTCCAGTTTTTACCCAAATGGGTTGGGGATCTCCCGAATGGTAGA  
GTGTGTCCCTGAGATGGAATCAGCTTGAGTCTTCTGCAATTGGTCACAACTATTCATGCT  
TCCTGTGATTTTCATCCAACCTACTTACCTTGCTACGATATCCCCTTTATCTCTAATCAGT  
TTATTTTCTTTCAAATAAAAAATAACTATGAGCAACATAAAAAAAAAAAAAA

197/392

## FIGURE 188

MKFLAVLVLLGVSI FLVSAQNPTTAAPADTYPATGPADDEAPDAETTAAATTATTAAPT  
TTAAASTTARKDIPVLPKWVGDL PNGRVCP

198/392

**FIGURE 189**

GAGCGAACATGGCAGCGCGTTGGCGGTTTTGGTGTGTCTCTGTGACCATGGTGGTGGCGC  
TGCTCATCGTTTTGCGACGTTCCCTCAGCCTCTGCCCAAAGAAAGAAGGAGATGGTGTAT  
CTGAAAAGGTTAGTCAGCTGATGGAATGGACTAACAAAAGACCTGTAATAAGAATGAATG  
GAGACAAGTTCGGTCGCCCTTGTGAAAGCCCCACCGAGAAATTACTCCGTTATCGTCATGT  
TCACTGCTCTCCAAGTGCATAGACAGTGTGTGCTTTGCAAGCAAGCTGATGAAGAATTCC  
AGATCCTGGCAAACTCCTGGCGATACTCCAGTGCATTACCAACAGGATATTTTTTGCCA  
TGGTGGATTTTGATGAAGGCTCTGATGTATTTAGATGCTAAACATGAATTCAGCTCCAA  
CTTTCATCAACTTTCTGCAAAAGGGAAACCCAAACGGGGTGATACATATGAGTTACAGG  
TGCGGGGTTTTTTCAGCTGAGCAGATTGCCCGGTGGATCGCCGACAGAACTGATGTCAATA  
TTAGAGTGATTAGACCCCAATTATGCTGGTCCCTTATGTTGGGATTGCTTTTGGCTG  
TTATTGGTGGACTTGTGTATCTTCGAAGAAGTAATATGGAATTTCTCTTTAATAAACTG  
GATGGGCTTTTGCAGCTTTGTGTTTTGTGCTTGCTATGACATCTGGTCAAATGTGGAACC  
ATATAAGAGGACCACCATATGCCATAAGAATCCCCACACGGGACATGTGAATTATATCC  
ATGGAAGCAGTCAAGCCAGTTTGTAGCTGAAACACACATTGTTCTTCTGTTTAAATGGTG  
GAGTTACCTTAGGAATGGTGCTTTTATGTGAAGCTGCTACCTCTGACATGGATATTGGAA  
AGCGAAAGATAATGTGTGTGGCTGGTATTGGACTTGTGTATTATTCTTCAGTTGGATGC  
TCTCTATTTTTAGATCTAAATATCATGGCTACCCATACAGCTTTCTGATGAGTTAAAAAG  
GTCCCAGAGATATATAGACACTGGAGTACTGGAAATTGAAAAACGAAAAATCGTGTGTGT  
TGAAAAGAAGAATGCAACTTGTATATTTTGTATTACCTCTTTTTTTCAAGTGATTTAAAT  
AGTTAATCATTTAACCAGAAAGATGTGTAGTGCCTTAACAAGCAATCCTCTGTCAAAAT  
CTGAGGTATTTGAAAATAATTATCCTCTTAACCTTCTCTTCCAGTGAACCTTATGGAAC  
ATTTAATTTAGTACAATTAAGTATATTATAAAAATTGTAAACTACTACTTTGTTTTAGT  
TAGAACAAAGCTCAAACTACTTTAGTTAAGTTGGTCATCTGATTTTATATTGCCTTATC  
CAAAGATGGGGAAAGTAAGTCCTGACCAGGTGTTCCCATATATGCCTGTACAGATAACT  
ACATTAGGAATTCATTCTTAGCTTCTTCATCTTTGTGTGGATGTGTATACTTTACGCATC  
TTTCTTTTGAGTAGAGAAATTATGTGTGTCATGTGGTCTTCTGAAAATGGAACACCATT  
CTTCAGAGCACACGTCTAGCCCTCAGCAAGACAGTTGTTTCTCCTCCTCCTTGCATATTT  
CCTACTGCGCTCCAGCCTGAGTGATAGAGTGAGACTCTGTCTCAAAAAAAGTATCTCTA  
AATACAGGATTATAATTTCTGCTTGAGTATGGTGTAACTACCTTGTATTTAGAAAGATT  
TCAGATTCATTCCATCTCCTTAGTTTTCTTTAAGGTGACCCATCTGTGATAAAAATATA  
GCTTAGTGCTAAAATCAGTGTAACCTTATACATGGCCTAAAATGTTTCTACAAATTAGAGT  
TTGTCACTTATTCATTTGTACCTAAGAGAAAAATAGGCTCAGTTAGAAAAGGACTCCCT  
GGCCAGGCGCAGTGACTTACGCCTGTAATCTCAGCACTTTGGGAGGCCAAGGCAGGCAGA  
TCACGAGGTCAGGAGTTCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAA  
AAATATAAAAATTAGCTGGGTGTGGTGGCAGGAGCCTGTAATCCCAGCTACACAGGAGGC  
TGAGGCACGAGAATCACTTGAACCTCAGGAGATGGAGGTTTCAGTGAGCCGAGATCACGCC  
ACTGCACTCCAGCCTGGCAACAGAGCGAGACTCCATCTCAAAAAAAAAAAAAA

199/392

**FIGURE 190**

MAARWRFWCVSVTMVVALLVCDVPSASAQRKKEMVLSEKVSQLMWETNKRFPVIRMNGDK  
FRRLVKAPPRNYSVIVMFTALQLHRQCVVCKQADEEFQILANSWRYSSAFTNRIFFAMVD  
FDEGSDVFQMLNMNSAPTFFINFPKAGKPKRGDTYELQVRGFSAEQIARWIADRTDVNIRV  
IRPPNYAGPLMLGLLLAVIGGLVYLRRSNMEFLFNKTGWAFALCFVLAMTSGQMWNHIR  
GPPYAHKNPHTGHVNYIHGSSQAQFVAETHIVLLFNNGVTLGMVLLCEAATSDMDIGKRK  
IMCVAGIGLVVLFFSWMLSIFRSKYHGYPYSFLMS

**Signal peptide:**  
amino acids 1-29

**Transmembrane domains:**  
amino acids 183-205, 217-237, 217-287, 301-321

200/392

**FIGURE 191**

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCA  
AGAGCTTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGC  
CATGGCCTCTCTTGGCCTCCAACTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGG  
CACACTGGTTGCCATGCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCAT  
TGTGACAGCAGTTGGCTTCTCCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGG  
CATCACCCAGTGTGACATCTATAGCACCTTCTGGGCCTGCCCGCTGACATCCAGGCTGC  
CCAGGCCATGATGGTGACATCCAGTGCAATCTCCTCCCTGGCCTGCATTATCTCTGTGGT  
GGGCATGAGATGCACAGTCTTCTGCCAGGAATCCCGAGCCAAAGACAGAGTGGCGGTAGC  
AGGTGGAGTCTTTTTCATCCTTGGAGGCCTCCTGGGATTCAATCCTGTTGCCTGGAATCT  
TCATGGGATCCTACGGGACTTCTACTCACCCTGGTGCCTGACAGCATGAAATTTGAGAT  
TGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAGCTGGAATCAT  
CCTCTGCTTTTCTGCTCATCCCAGAGAAATCGCTCCAACCTACTACGATGCCTACCAAGC  
CCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA  
GTTCAATTCTACAGCCTGACAGGGTATGTGTGAAAGAACCAGGGGCCAGAGCTGGGGGGT  
GGCTGGGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACC  
ACTGGATCGTGTGAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCA  
AAGGCAGAAATGGGGGCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCA  
TGCCAGCCTTTCTGTTTTCTCACCTTGCTGCTCCCCTGCCCTAAGTCCCCAACCCCTCAA  
CTTGAAACCCCATTCCTTAAGCCAGGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCT  
GGGACTCCATCCCCAAACCCACTAATCACATCCCCTGACTGACCCTCTGTGATCAAAGA  
CCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATTGCTGGGGATGGGAAGGAGAAGCAGT  
GGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTCCCTCCAAAGAACTGATTGG  
CCCTGGAACCTCCATCCCCTCTTGTATGACTCCACAGTGTCCAGACTAATTTGTGCAT  
GAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATGCAGGATGGGA  
GGACAGGAAGGCAGCCTGGGACATTTAAAAAATA



201/392

## FIGURE 192

MASLGLQLVGYILGLLGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTG  
ITQCDIYSTLLGLPADIQAAQAMMTSSAIISSSLACIIISVVGMRCTVFCQESRAKDRVAVA  
GGVFFILGGLLGFIIPVAVNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSSLFSLIAGII  
LCFSCSSQRNRSNYYDAYQAQPLATRSSPRPGQPPKVKSEFNSSSLTGYV

**Important features of the protein:**

**Signal peptide:**

amino acids 1-24

**Transmembrane domains:**

amino acids 82-102, 117-140, 163-182

**N-glycosylation site:**

amino acids 190-193

**PMP-22 / EMP / MP20 family proteins:**

amino acids 46-59

202/392

**FIGURE 193**

CTCCACTGCAACCACCCAGAGCCATGGCTCCCCGAGGCTGCATCGTAGCTGTCTTTGCCA  
TTTTCTGCATCTCCAGGCTCCTCTGCTCACACGGAGCCCCAGTGGCCCCCATGACTCCTT  
ACCTGATGCTGTGCCAGCCACACAAGAGATGTGGGGACAAGTTCTACGACCCCCCTGCAGC  
ACTGTTGCTATGATGATGCCGTCGTGCCCTTGGCCAGGACCCAGACGTGTGGAACTGCA  
CCTTCAGAGTCTGCTTTGAGCAGTGCTGCCCCCTGGACCTTCATGGTGAAGCTGATAAACC  
AGAACTGCGACTCAGCCCGGACCTCGGATGACAGGCTTTGTGCGCAGTGTCAGCTAATGGA  
ACATCAGGGGAACGATGACTCCTGGATTCTCCTTCCTGGGTGGGCCTGGAGAAAGAGGCT  
GGTGTTACCTGAGATCTGGGATGCTGAGTGGCTGTTTGGGGGCCAGAGAAAACACACTC  
AACTGCCCACTTCATTCTGTGACCTGTCTGAGGCCCACCCTGCAGCTGCCCTGAGGAGGC  
CCACAGGTCCCCCTTCTAGAATTCTGGACAGCATGAGATGCCGTGTGCTGATGGGGGCCAG  
GGACTCTGAACCCTCCTGATGACCCCTATGGCCAACATCAACCCGGCACCACCCCAAGGC  
TGGCTGGGGAACCCCTCACCCCTTCTGTGAGATTTTCCATCATCTCAAGTTCTCTTCTATC  
CAGGAGCAAAGCACAGGATCATAATAAATTTATGTACTTTATAAATGAAA

203/392

## FIGURE 194

MAPRGCIVAVFAIFCISRLLC SHGAPVAPMTPYLMLCQPHKRCGDKFYDPLQHCCYDDAV  
VPLARTQTCGNCTFRVCFEQCCPWTFMVKLINQNCD SARTSDDR LCRSVS

Signal peptide:  
amino acids 1-24

204/392

**FIGURE 195**

CATTTCCAACAAGAGCACTGGCCAAGTCAGCTTCTTCTGAGAGAGTCTCTAGAAGACATG  
ATGCTACACTCAGCTTTGGGTCTCTGCCTCTTACTCGTCACAGTTTCTTCCAACCTTGCC  
ATTGCAATAAAAAAGGAAAAGAGGCCTCCTCAGACACTCTCAAGAGGATGGGGAGATGAC  
ATCACTTGGGTACAACTTATGAAGAAGGTCTCTTTTATGCTCAAAAAAGTAAGAAGCCA  
TTAATGGTTATTTCATCACCTGGAGGATTGTCAATACTCTCAAGCACTAAAGAAAGTATTT  
GCCCCAAATGAAGAAATACAAGAAATGGCTCAGAATAAGTTCATCATGCTAAACCTTATG  
CATGAAACCACTGATAAGAATTTATCACCTGATGGGCAATATGTGCCTAGAATCATGTTT  
GTAGACCCTTCTTTAACAGTTAGAGCTGACATAGCTGGAAGATACTCTAACAGATTGTAC  
ACATATGAGCCTCGGGATTTACCCCTATTGATAGAAAACATGAAGAAAGCATTAAAGACTT  
ATTCAGTCAGAGCTATAAGAGATGATGGAAAAAGCCTTCACCTCAAAGAAGTCAAATTT  
CATGAAGAAAACCTCTGGCACATTGACAAATACTAAATGTGCAAGTATATAGATTTTGTA  
ATATTACTATTTAGTTTTTTTAAATGTGTTTGCAATAGTCTTATTAAAATAAATGTTTTTT  
AAATCTGA

205/392

## FIGURE 196

MMLHSALGLCLLLVTVSSNLAIKKEKRPPQTL SRGWGDDITWVQTYEEGLFYAQSKK  
PLMVIHHLED CQYSQALKKVFAQNEEIQEMAQNKFIMLNLMHETTDKNLSPDGQYVPRIM  
FVDPSLTVRADIAGRYSNRLYTYEPRDLPLLIENMKKALRLIQSEL

**Important features:**

**Signal peptide:**

amino acids 1-23

**N-myristoylation site:**

amino acids 51-57

206/392

**FIGURE 197**

GGGGGCGGGTGCCTGGAGCACGGCGCTGGGGCCGCCCCGAGCGCTCACTCGCTCGCACTC  
AGTCGCGGGAGGCTTCCCCGCGCCGGCCGCGTCCCCGCCGCTCCCCGGCACCAGAAGTTC  
CTCTGCGCGTCCGACGGCGACATGGGCGTCCCCACGGCCCTGGAGGCCGGCAGCTGGCGC  
TGGGGATCCCTGCTCTTCGCTCTCTTCTGGCTGCGTCCCTAGGTCCGGTGGCAGCCTTC  
AAGGTCGCCACGCCGTATTCCCTGTATGTCTGTCCCGAGGGGCAGAACGTCAACCTCACC  
TGCAGGCTCTTGGGGCCCTGTGGACAAAGGGCACGATGTGACCTTCTACAAGACGTGGTAC  
CGCAGCTCGAGGGGCGAGGTGCAGACCTGCTCAGAGCGCCGGCCCATCCGCAACCTCACG  
TTCCAGGACCTTCACCTGCACCATGGAGGCCACCAGGCTGCCAACACCAGCCACGACCTG  
GCTCAGCGCCACGGGCTGGAGTCGGCCTCCGACCACCATGGCAACTTCTCCATCACCATG  
CGCAACCTGACCCCTGCTGGATAGCGGCCTCTACTGCTGCCTGGTGGTGGAGATCAGGCAC  
CACCATCCGAGCACAGGGTCCATGGTGCATGGAGCTGCAGGTGCAGACAGGCAAAGAT  
GCACCATCCAACGTGTGTGGTGTACCCATCCTCCTCCCAGGATAGTGAAAACATCACGGCT  
GCAGCCCTGGCTACGGGTGCCTGCATCGTAGGAATCCTCTGCCTCCCCCTCATCTGCTC  
CTGGTCTACAAGCAAAGGCAGGCAGCCTCCAACCGCCGTGCCAGGAGCTGGTGC GGATG  
GACAGCAACATTCAAGGGATTGAAAACCCCGGCTTTGAAGCCTCACCACCTGCCCAGGGG  
ATACCCGAGGCCAAAGTCAGGCACCCCTGTCTTATGTGGCCAGCGGCAGCCTTCTGAG  
TCTGGGCGGCATCTGCTTTCGGAGCCCAGCACCCCTGTCTCCTCCAGGCCCGGAGAC  
GTCTTCTTCCCATCCCTGGACCTGTCCCTGACTCTCCAAACTTTGAGGTCTCTAGCCC  
AGCTGGGGGACAGTGGGCTGTTGTGGCTGGGTCTGGGGCAGGTGCATTTGAGCCAGGGCT  
GGCTCTGTGAGTGGCCTCCTTGGCCTCGGCCCTGGTTCCTCCTCCTGCTCTGGGCTCA  
GATACTGTGACATCCAGAACCCAGCCCTCAACCCCTCTGGATGCTACATGGGGATGC  
TGGACGGCTCAGCCCCGTGTTCCAAGGATTTTGGGGTGCTGAGATTCTCCCCTAGAGACCT  
GAAATTCACCAGCTACAGATGCCAAATGACTTACATCTTAAGAAGTCTCAGAACGTCCAG  
CCCTTCAGCAGCTCTCGTTCTGAGACATGAGCCTTGGGATGTGGCAGCATCAGTGGGACA  
AGATGGACACTGGGGCCACCCTCCAGGCACCAGACACAGGGCACGGTGGAGAGACTTCTC  
CCCCGTGGCCGCCTTGGCTCCCCCGTTTTGGCCGAGGCTGCTCTTCTGTGAGACTTCTC  
TTTGTACCACAGTGGCTCTGGGGCCAGGCCTGCCTGCCCCTGGCCATCGCCACCTTCCC  
CAGCTGCCTCCTACCAGCAGTTTCTCTGAAGATCTGTCAACAGGTTAAGTCAATCTGGGG  
CTTCCACTGCCTGCATTCCAGTCCCCAGAGCTTGGTGGTCCCGAAACGGGAAGTACATAT  
TGGGGCATGGTGGCCTCCGTGAGCAATGGTGTCTTGGGCAATCTGAGGCCAGGACAGAT  
GTTGCCCCACCCACTGGAGATGGTGTCTGAGGGAGGTGGGTGGGGCCTTCTGGGAAGGTGA  
GTGGAGAGGGGCACCTGCCCCCGCCCTCCCCATCCCTACTCCCACTGCTCAGCGCGGG  
CCATTGCAAGGGTGCCACACAATGTCTTGTCCACCCCTGGGACACTTCTGAGTATGAAGCG  
GGATGCTATTAAAACTACATGGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAGA

207/392

## FIGURE 198

MGVPTALEAGSWRWGSLFLAASLGPVAAFQVATPYSLYVCPEGQNVTLTCRLLGPV  
DKGHDVTFYKTYWRSSRGEVQTCSERRPIRNLTFQDLHLHHGGHQAANTSHDLAQRHGLE  
SASDHHGNFSITMRNLTLDSGLYCCLVVEIRHHHSEHRVHGAMELQVQTGKDAPSNCVV  
YPSSSQDSENITAAALATGACIVGILCLPLILLLVYKQRQAASNRRQELVRMDSNIQGI  
ENPGFEASPPAQGIPEAKVRHPLSYVAQRQPSESGRHLLSEPSTPLSPPGPDVFFPSLD  
PVPDSPNFEVI

**Signal peptide:**  
amino acids 1-28

**Transmembrane domain:**  
amino acids 190-216

208/392

**FIGURE 199**

CTAGCCTGCGCCAAGGGGTAGTGAGACCGCGCGGCAACAGCTTGCGGCTGCGGGGAGCTC  
CCGTGGGCGCTCCGCTGGCTGTGCAGGCGGCCATGGATTCTTGCGGAAAATGCTGATCT  
CAGTCGCAATGCTGGGCGCAGGGGCTGGCGTGGGCTACGCGCTCCTCGTTATCGTGACCC  
CGGGAGAGCGGCGGAAGCAGGAAATGCTAAAGGAGATGCCACTGCAGGACCCAAGGAGCA  
GGGAGGAGGCGGCCAGGACCCAGCAGCTATTGCTGGCCACTCTGCAGGAGGCAGCGACCA  
CGCAGGAGAACGTGGCCTGGAGGAAGAACTGGATGGTTGGCGGCGAAGGCGGCGCCAGCG  
GGAGGTCACCGTGAAGACCGGACTTGCCTCCGTGGGCGCCGGACCTTGGCTTGGGCGCAGG  
AATCCGAGGCAGCCTTTCTCCTTCGTGGGCCAGCGGAGAGTCCGGACCGAGATACCATG  
CCAGGACTCTCCGGGGTCCTGTGAGCTGCCGTCCGGTGAGCACGTTTCCCCCAAACCCTG  
GACTGACTGCTTTAAGGTCCGCAAGGCGGGCCAGGGCCGAGACGCGAGTCGGATGTGGTG  
AACTGAAAGAACCAATAAAATCATGTTCTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA



209/392

## FIGURE 200

MDSLRKMLISVAMLGAGAGVGYALLVIVTPGERRKQEMLKEMPLQDPRSREEAARTQQLL  
LATLQEAATTQENVAWRKNNMVGEGGASGRSP

Signal peptide:  
amino acids 1-18

210/392

**FIGURE 201**

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGGCCCTCCGCTCACGC  
AGAGCCTCTCCGTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTCTAAT  
CCATCCGTCACCTCTCCTGTTCATCCGTTTCCATGCCGTGAGGTCCATTACAGAACACAT  
CCATGGCTCTCATGCTCAGTTTGGTTCTGAGTCTCCTCAAGCTGGGATCAGGGCAGTGGC  
AGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCCTTGGTGGGGGAGGACGCAGCATTCTCCT  
GTTTCCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGGCCAGT  
TCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCAC  
AGTATCAAGGCAGGACAAAACCTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGA  
GGCTGGAAAACATTACTGTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCAGT  
CTTACTACCAGAAGGCCATCTGGGAGCTACAGGTGTCAGCACTGGGCTCAGTTCTCTCA  
TTTCCATCACGGGATATGTTGATAGAGACATCCAGCTACTCTGTCACTCCTCGGGCTGGT  
TCCCCCGGGCCACAGCGAAGTGGAAGGTCCACAAGGACAGGATTTGTCCACAGATCCA  
GGACAAACAGAGACATGCATGGCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGA  
ACGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCA  
GGGTACAGATAGGAGATACCTTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTAC  
TGGAATACTCTGCTGTGGCCTATTTTTTGGCATTGTTGGACTGAAGATTTTTCTCTCCA  
AATTCAGTGGAATAATCCAGGCGGAACCTGGACTGGAGAAGAAAGCACGGACAGGCAGAAT  
TGAGAGACGCCCCGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGA  
AGCTCTGCGTTTCTGATCTGAAAACCTGTAACCCATAGAAAAGCTCCCAGGAGGTGCCTC  
ACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGGA  
AACATTACTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCGTGGGAGTGTGCCGGG  
ATGATGTGGACAGGAGGAAGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCC  
TCAGACTGAATGGAGAACATTTGTATTTTACATTAAATCCCCGTTTTATCAGCGTCTTCC  
CCAGGACCCACCTACAAAAATAGGGGTCTTCTGGACTATGAGTGTGGGACCATCTCCT  
TCTTCAACATAAATGACCAGTCCCTTATTTATACCTGACATGTCGGTTTGAAGGCTTAT  
TGAGGCCCTACATTGAGTATCCGTCTTATAATGAGCAAAATGGAACCTCCCATAGTCATCT  
GCCCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGCAATCCCAG  
AGACAAGCAACAGTGAGTCTCTCTCACAGGCAACCACGCCCTTCTCCCCAGGGGTGAAA  
TGTAAGGATGAATCACATCCCACATTCTTCTTTAGGGATATTAAGGTCTCTCTCCAGATC  
CAAAGTCCCGCAGCAGCCGGCCAAGGTGGCTTCCAGATGAAGGGGGACTGGCCTGTCCAC  
ATGGGAGTCAGGTGTCTATGGCTGCCCTGAGCTGGGAGGGAAGAAGGCTGACATTACATTT  
AGTTTGCTCTCACTCCATCTGGCTAAGTGATCTTGAAATACCACCTCTCAGGTGAAGAAC  
CGTCAGGAATTCCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATA  
ATGCTTAGATCTTATTGATGACAGAGTGATCCTAATGGTTTGTTCATTATATTACACTT  
TCAGTAAAAAAA

211/392

**FIGURE 202**

MALMLSLVLSLLKLGSQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQF  
SSVVHLYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISQS  
YYQKAIWELQVSALGSVPLISITGYVDRDIQLLCQSSGWFFRPTAKWKGPQGQDLSTDSR  
TNRDMHGLFDVEISLTVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVL  
GILCCGLFFGIVGLKIFFSKFQWKIQAEILDWRRKHGQAEIRDARKHAVEVTLDPETAHPK  
LCVSDLKTVTHRKAPEVPHSEKRFTRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVCRD  
DVDRRKEYVTLSPDHGYWVLRNLNGEHLTYFTLNPRFISVFPRTPTKIGVFLDYECGTISF  
FNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGTPIVICPVTQESEKEASWQRASAIPE  
TSNSESSSQATTPLPRGEM

**Signal peptide:**  
amino acids 1-17

**Transmembrane domain:**  
amino acids 239-255

212/392

**FIGURE 203**

TGCGGCGCAGTGCTAGACCTGGGAGGATGGGCGGCCTGCTGCTGGCTGCTTTTCTGGCTTT  
GGTCTCGGTGCCCAGGGCCCAGGCCGTGTGGTTGGGAAGACTGGACCCTGAGCAGCTTCT  
TGGGCCCTGGTACGTGCTTGCGGTGGCCTCCCGGAAAAGGGCTTTGCCATGGAGAAGGA  
CATGAAGAACGTCGTGGGGGTGGTGGTGACCCTCACTCCAGAAAACAACCTGCGGACGCT  
GTCCTCTCAGCACGGGCTGGGAGGGTGTGACCAGAGTGTCATGGACCTGATAAAGCGAAA  
CTCCGATGGGTGTTTGAGAATCCCTCAATAGGCGTGCTGGAGCTCTGGGTGCTGGCCAC  
CAACTTCAGAGACTATGCCATCATCTTCACTCAGCTGGAGTTCGGGGACGAGCCCTTCAA  
CACCGTGGAGCTGTACAGTCTGACGGAGACAGCCAGCCAGGAGGCCATGGGGCTCTTCAC  
CAAGTGGAGCAGGAGCCTGGGCTTCTGTGTCACAGTAGCAGGCCCAGCTGCAGAAGGACCT  
CACCTGTGCTCACAAGATCCTTCTGTGAGTGCTGCGTCCCCAGTAGGGATGGCGCCCACA  
GGGTCTGTGACCTCGGCCAGTGTCACCCACCTCGCTCAGCGGCTCCCGGGGCCAGCA  
CCAGCTCAGAATAAAGCGATTCCACAGCA

213/392

## FIGURE 204

MGGLLLAAFLALVSVPRQAVWLGRDLPEQLLGPWYVLAVASREKGFAMEKDMKNVVGVV  
VTLTPENNLRTLSSQHGLGGCDQSVMDLIKRNSGWVFENPSIGVLELWVLATNFRDYAI I  
FTQLEFGDEPFNTVELYSITETASQEAMGLFTKWSRSLGFLSQ

Signal peptide:  
amino acids 1-20

214/392

**FIGURE 205**

GACGCCCAGTGACCTGCCGAGGTCGGCAGCACAGAGCTCTGGAGATGAAGACCCTGTTCC  
TGGGTGTCACGCTCGGCCTGGCCGCTGCCCTGTCCTTCACCCTGGAGGAGGAGGATATCA  
CAGGGACCTGGTACGTGAAGGCCATGGTGGTCGATAAGGACTTTCCGGAGGACAGGAGGC  
CCAGGAAGGTGTCCCCAGTGAAGGTGACAGCCCTGGGCGGTGGGAAGTTGGAAGCCACGT  
TCACCTTCATGAGGGAGGATCGGTGCATCCAGAAGAAAATCCTGATGCGGAAGACGGAGG  
AGCCTGGCAAATACAGCGCCTATGGGGGCAGGAAGCTCATGTACCTGCAGGAGCTGCCCCA  
GGAGGGACCACTACATCTTTTACTGCAAAGACCAGCACCATGGGGGCCTGCTCCACATGG  
GAAAGCTTGTGGGTAGGAATTCTGATACCAACCGGGAGGCCCTGGAAGAATTTAAGAAAT  
TGGTGCAGCGCAAGGGACTCTCGGAGGAGGACATTTTCACGCCCCTGCAGACGGGAAGCT  
GCGTTCCCGAACACTAGGCAGCCCCCGGGTCTGCACCTCCAGAGCCCACCCTACCACCAG  
ACACAGAGCCCCGGACCACCTGGACCTACCCTCCAGCCATGACCCTTCCCTGCTCCCACCC  
ACCTGACTCCAAATAAAGTCCTTTTCCCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAA

215/392

## FIGURE 206

MKTLFLGVTLGLAAALSFTLEEDITGTWYVKAMVVDKDFPEDRRPRKVSPVKVTALGGG  
KLEATFTFMREDRCIQKKILMRKTEEPGKYSAYGGRKLMYLQELPRRDHYIFYCKDQHHG  
GLLHMGKLVGRNSDTNREALEEFKKLVQRKGLSEEDIFTPLQTGSCVPEH

Important features:

Signal peptide:

amino acids 1-17

216/392

**FIGURE 207**

GTTCCGCAGATGCAGAGGTTGAGGTGGCTGCGGGACTGGAAGTCATCGGGCAGAGGTCTC  
ACAGCAGCCAAGGAACCTGGGGCCCGCTCCTCCCCCTCCAGGCCATGAGGATTCTGCAG  
TTAATCCTGCTTGCTCTGGCAACAGGGCTTGTTAGGGGGAGAGACCAGGATCATCAAGGGG  
TTCGAGTGCAAGCCTCACTCCCAGCCCTGGCAGGCAGCCCTGTTTCGAGAAGACGCGGCTA  
CTCTGTGGGGCGACGCTCATCGCCCCAGATGGCTCCTGACAGCAGCCCACTGCCTCAAG  
CCCCGCTACATAGTTCACCTGGGGCAGCACAACTCCAGAAGGAGGAGGGCTGTGAGCAG  
ACCCGGACAGCCACTGAGTCCTTCCCCACCCCGGCTTCAACAACAGCCTCCCCAACAAA  
GACCACCGCAATGACATCATGCTGGTGAAGATGGCATCGCCAGTCTCCATCACCTGGGCT  
GTGCGACCCCTCACCTCTCCTCACGCTGTGTCACTGCTGGCACCAGCTGCCTCATTTCC  
GGCTGGGGCAGCACGTCCAGCCCCAGTTACGCCTGCCTCACACCTTGCGATGCGCCAAC  
ATCACCATCATTGAGCACCAGAAGTGTGAGAACGCCTACCCCGGCAACATCACAGACACC  
ATGGTGTGTGCCAGCGTGCAGGAAGGGGGCAAGGACTCCTGCCAGGGTGACTCCGGGGGC  
CCTCTGGTCTGTAAACAGTCTCTTCAAGGCATTATCTCCTGGGGCCAGGATCCGTGTGCG  
ATCACCCGAAAGCCTGGTGTCTACACGAAAGTCTGCAAATATGTGGACTGGATCCAGGAG  
ACGATGAAGAACAATTAGACTGGACCCACCCACCACAGCCCATCACCTCCATTTCCACT  
TGGTGTGTTGGTTCTGTTCACTCTGTTAATAAGAAACCCTAAGCCAAGACCTCTACGAA  
CATTCTTTGGGCCTCCTGGACTACAGGAGATGCTGTCACTTAATAATCAACCTGGGGTTC  
GAAATCAGTGAGACCTGGATTCAAATTCTGCCTTGAAATATTGTGACTCTGGGAATGACA  
ACACCTGGTTTGTCTCTGTTGTATCCCCAGCCCCAAAGACAGCTCCTGGCCATATATCA  
AGGTTTCAATAAATATTTGCTAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAA



217/392

**FIGURE 208**

MRILQLILLALATGLVGGETRIIKGFECKPHSQPWQAALFEKTRLLCGATLIAPRWLLTA  
AHCLKPRYIVHLGQHNLQKEEGCEQTRTATESFPHPGFNNSLPNKDHRNDIMLVKMASPV  
SITWAVRPLTLSSRCVTAGTSCLISGWGSTSSPQLRLPHTLRCANITIIHQKCENAYPG  
NITDTMVCASVQEGGKDSCQGDSSGGLVCNQSLQGIISWGQDPCAITRKPGVYTKVCKYV  
DWIQETMKNN

**Important features:****Signal peptide:**

amino acids 1-18

**Serine proteases, trypsin family, histidine active site:**

amino acids 58-63

**N-glycosylation sites:**

amino acids 99-102, 165-168, 181-184, 210-213

**Glycosaminoglycan attachment site:**

amino acids 145-148

**Kringle domain proteins:**

amino acids 197-209, 47-64

**Serine proteases, trypsin family, histidine protein:**

amino acids 199-209, 47-63, 220-243

**Apple domain proteins:**

amino acids 222-249, 189-222

218/392

**FIGURE 209**

GCGGCCACACGCAGCTAGCCGGAGCCCGGACCAGGCGCCTGTGCCTCCTCCTCGTCCCTC  
GCCGCGTCCGCGAAGCCTGGAGCCGGCGGGAGCCCCGCGCTCGCCATGTCGGGCGAGCTC  
AGCAACAGGTTCCAAGGAGGGAAGGCGTTCGGCTTGCTCAAAGCCCGGCAGGAGAGGAGG  
CTGGCCGAGATCAACCGGGAGTTTCTGTGTGACCAGAAGTACAGTGATGAAGAGAACCTT  
CCAGAAAAGCTCACAGCCTTCAAAGAGAAGTACATGGAGTTTGACCTGAACAATGAAGGC  
GAGATTGACCTGATGTCTTTAAAGAGGATGATGGAGAAGCTTGGTGTCCCAAGACCCAC  
CTGGAGATGAAGAAGATGATCTCAGAGGTGACAGGAGGGGTGTCAGTGACACTATATCCTAC  
CGAGACTTTGTGAACATGATGCTGGGGAAACGGTCGGCTGTCCTCAAGTTAGTCATGATG  
TTTGAAGGAAAAGCCAACGAGAGCAGCCCCAAGCCAGTTGGCCCCCTCCAGAGAGAGAC  
ATTGCTAGCCTGCCCTGAGGACCCCGCCTGGACTCCCAGCCTTCCCACCCCATACCTCC  
CTCCCGATCTTGCTGCCCTTCTTGACACACTGTGATCTCTCTCTCTCTCATTGTGTTGGT  
CATTGAGGGTTTGTGTTGTGTTTTTCATCAATGTCTTTGTAAAGCACAAATTATCTGCCTTA  
AAGGGGCTCTGGGTGCGGGAATCCTGAGCCTTGGGTCCCCTCCCTCTCTTCTTCCCTCCT  
TCCCCGCTCCCTGTGCAGAAGGGCTGATATCAAACCAAAAAGTAGAGGGGGCAGGGCCAG  
GGCAGGGAGGCTTCCAGCCTGTGTTCCCCTCACTTGGAGGAACCAGCACTCTCCATCCTT  
TCAGAAAGTCTCCAAGCCAAGTTCAGGCTCACTGACCTGGCTCTGACGAGGACCCCAGGC  
CACTCTGAGAAGACCTTGGAGTAGGGACAAGGCTGCAGGGCCTCTTTCGGGTTTCCTTGG  
ACAGTGCCATGGTTCCAGTGCTCTGGTGTCAACCAGGACACAGCCACTCGGGGCCCCGCT  
GCCCCAGCTGATCCCCACTCATTCCACACCTCTTCTCATCCTCAGTGATGTGAAGGTGGG  
AAGGAAAGGAGCTTGGCATTGGGAGCCCTTCAAGAAGGTACCAGAAGGAACCCTCCAGTC  
CTGCTCTCTGGCCACACCTGTGCAGGCAGCTGAGAGGCAGCGTGCAGCCCTACTGTCCCT  
TACTGGGGCAGCAGAGGGCTTCGGAGGCAGAAGTGAGGCCTGGGGTTTGGGGGGAAAGGT  
CAGCTCAGTGCTGTTCCACCTTTTAGGGAGGATACTGAGGGGACCAGGATGGGAGAATGA  
GGAGTAAAATGCTCACGGCAAAGTCAGCAGCACTGGTAAGCCAAGACTGAGAAATACAAG  
GTTGCTTGTCTGACCCCAATCTGCTTGAAAAAAAAAAAAAAAAAAAAA

219/392

## FIGURE 210

MSGELSNRFQGGKAFGLLKARQERRLAELNREFLCDQKYSDEENLPEKLTAFKEKYMED  
LNNEGEIDLMSLKRMMEKLGVPKTHLEMKKMISEVTGGVSDTISYRDFVNMMLGKRSAVL  
KLVMMFEGKANESSPKPVGPPPERDIASLP

220/392

**FIGURE 211**

CTGGGATCAGCCACTGCAGCTCCCTGAGCACTCTCTACAGAGACGCGGACCCAGACATG  
AGGAGGCTCCTCCTGGTCACCAGCCTGGTGGTTGTGCTGCTGTGGGAGGCAGGTGCAGTC  
CCAGCACCCAAGGTCCCTATCAAGATGCAAGTCAAACACTGGCCCTCAGAGCAGGACCCA  
GAGAAGGCCTGGGGCGCCCGTGTGGTGGAGCCTCCGGAGAAGGACGACCAGCTGGTGGTG  
CTGTTCCCTGTCCAGAAGCCGAAACTCTTGACCACCGAGGAGAAGCCACGAGGTCAGGGC  
AGGGGCCCCATCCTTCCAGGCACCAAGGCCTGGATGGAGACCGAGGACACCCTGGGCCGT  
GTCCTGAGTCCCGAGCCCGACCATGACAGCCTGTACCACCCTCCGCCTGAGGAGGACCAG  
GGCGAGGAGAGGCCCCGGTTGTGGGTGATGCCAAATCACCAGGTGCTCCTGGGACCGGAG  
GAAGACCAAGACCACATCTACCACCCCCAGTAGGGCTCCAGGGGCCATCACTGCCCCCGC  
CCTGTCCCAAGGCCCAGGCTGTTGGGACTGGGACCCCTCCCTACCCTGCCCCAGCTAGACA  
AATAAACCCAGCAGGCAAAAAAAAAAAAAAAAAAAAA

221/392

## FIGURE 212

MRRLLLVTSLVVVLLWEAGAVPAPKVPIKMQVKHWPSEQDPEKAWGARVVEPPEKDDQLV  
VLFPVQKPKLLTTEEKPRGQGRGPILPGTKAWMETEDTLGRVLSPEPDHDSLYHPPPEED  
QGEERPRLWVMPNHQVLLGPPEEDQDHIYHPQ

222/392

**FIGURE 213**

CAGGCAGAAGCGAACAAAGACCCAGCAAGAGAAGGCAGAGGCTAAGACCCATCCCGTATC  
TGCTCTCCTGAAATAATTCTGGAGTCATGCCTGAAATGCCAGAGGACATGGAGCAGGAGG  
AAGTTAACATCCCTAATAGGAGGGTTCTGGTTACTGGTGCCACTGGGCTTCTTGGCAGAG  
CTGTACACAAAGAAATTTAGCAGAATAATTGGCATGCAGTTGGCTGTGGTTTCAGAAGAG  
CAAGACCAAAATTTGAACAGGTTAATCTGTTGGATTCTAATGCAGTTCATCACATCATTC  
ATGATTTTCAGCCCCATGTTATAGTACATTGTGCAGCAGAGAGAAGACCAGATGTTGTAG  
AAAATCAGCCAGATGCTGCCTCTCAACTTAATGTGGATGCTTCTGGGAATTTAGCAAAGG  
AAGCAGCTGCTGTTGGAGCATTCTCATCTACATTAGCTCAGATTATGTATTTGATGGAA  
CAAATCCACCTTACAGAGAGGAAGACATACCAGCTCCCCTAAATTTGTATGGCAAAACAA  
AATTAGATGGAGAAAAGGCTGTCTGGAGAACAACTAGGAGCTGCTGTTTGGAGGATTC  
TATTCTGTATGGGGAAGTTGAAAAGCTCGAAGAAAGTGCTGTGACTGTTATGTTTGATA  
AAGTGCAGTTCAGCAACAAGTCAGCAAAACATGGATCACTGGCAGCAGAGGTTCCCCACAC  
ATGTCAAAGATGTGGCCACTGTGTGCCGAGCTAGCAGAGAAGAGAATGCTGGATCCAT  
CAATTAAGGGAACCTTTCACTGGTCTGGCAATGAACAGATGACTAAGTATGAAATGGCAT  
GTGCAATTGCAGATGCCTTCAACCTCCCCAGCAGTCACTTAAGACCTATTACTGACAGCC  
CTGTCTTAGGAGCACACGTCGAGAAATGCTCAGCTTGACTGCTCCAAATGGAGACCT  
TGGGCATTGGCCAACGAACACCATTTTGAATTGGAATCAAAGAATCACTTTGGCCTTTCC  
TCATTGACAAGAGATGGAGACAAACGGTCTTTTCACTTAGTTTATTTGTGTTGGGTTCTTTT  
TTTTTTTAAATGAAAAGTATAGTATGTGGCACTTTTTTAAAGAACAAAGGAAATAGTTTTG  
TATGAGTACTTTAATTGTGACTCTTAGGATCTTTCAGGTAAATGATGCTCTTGCACTAGT  
GAAATTGTCTAAAGAACTAAAGGGCAGTCATGCCCTGTTTGCAGTAATTTTCTTTTAA  
TCATTTTGTGTTGTCCTGGCTAAACTTGGAGTTTGAGTATAGTAAATTATGATCCTTAAAT  
ATTTGAGAGTCAGGATGAAGCAGATCTGCTGTAGACTTTTCAGATGAAATTGTTCAATTCT  
CGTAACCTCCATATTTTCAGGATTTTGAAGCTGTTGACCTTTTCATGTTGATTATTTTA  
AATTGTGTGAAATAGTATAAAAATCATTTGGTGTTCATTATTTGCTTTGCCTGAGCTCAGA  
TCAAAATGTTTGAAGAAAGGAACCTTTATTTTGAAGTTACGTACAGTTTTTATGCTTGA  
GATATTTCAACATGTTATGTATATTGGAACCTTCTACAGCTTGATGCCTCCTGCTTTTATA  
GCAGTTTATGGGGAGCACCTTGAAGAGCGTGTGTACATGTATTTTTTTCTAGGCAACA  
TTGAATGCAAACGTGTATTTTTTTAATATAAATATATAACTGTCCTTTTCATCCCATGTT  
GCCGCTAAGTGATATTTTATATGTGTGGTTATACTCATAAATGGGCCTTGTAAGTCTT  
TTCACCATTCATGAATAATAAATAATATGTACTGCTGGCATGTAATGCTTAGTTTCTTG  
TATTTACTTCTTTTTTTTAAATGTAAGGACCAACTTCTAACTAATTGTTCTTTGTTGC  
TTTAATTTTAAAAATTACATTCTTCTGATGTAACATGTGATACATACAAAAGAATATAG  
TTTAATATGTATTGAAATAAAACACAATAAAAT

223/392

**FIGURE 214**

MPEMPEDMEQEEVNIPNRRVLVTGATGLLGRAVHKEFQQNNWHAVGCGFRRARPKFEQVN  
LLDSNAVHHI IHDFQPHVIVHCAAERRPDVVENQPDAAASQLNVDASGNLAKEAAAVGAFL  
IYISSDYVFDGTNPPYREEDI PAPLNLYGKTKLDGEKAVLENNLGAAVLRIPILYGEVEK  
LEESAVTVMFDKVQFSNKSANMDHWQQRFP THVKDVATVCRQLAEKRMLDPSIKGTFHWS  
GNEQMTKYEMACAIADAFNLPSSHLRPITDSPVLGAQRPRNAQLDCSKLETLGIGQRTPF  
RIGIKESLWPFLIDKRWRQTVFH

**Signal peptide:**  
amino acids 1-30

**Transmembrane domain:**  
amino acids 105-127

**N-glycosylation site:**  
amino acids 197-201

**N-myristoylation site:**  
amino acids 303-309

**Short-chain dehydrogenases/reductases family proteins:**  
amino acids 18-30

224/392

**FIGURE 215**

GTGAATGTGAGGGTTTGATGACTTTCAGATGTCTAGGAACCAGAGTGGGTGCAGGGGCCCC  
CAGGCAGGGCTGATTCTTGGGCGGAGGAGAGTAGGGTAAAGGGTTCTGCATGAGCTCCTT  
AAAGGACAAAGGTAACAGAGCCAGCGAGAGAGCTCGAGGGGAGACTTTGACTTCAAGCCA  
CAGAATTGGTGGAAGTGTGCGCGCCGCCGCCGCTCGCTCCTGCAGCGCTGTGACCTA  
GCCGCTAGCATCTTCCCGAGCACCGGGATCCCGGGGTAGGAGGCGACGCGGGCGAGCACC  
AGCGCCAGCCGGCTGCGGCTGCCCACACGGCTCACCATGGGCTCCGGGCGCCGGGCGCTG  
TCCGCGGTGCCGGCCGTGCTGCTGGTCCTCACGCTGCCGGGGCTGCCCGTCTGGGCACAG  
AACGACACGGAGCCCATCGTGCTGGAGGGCAAGTGTCTGGTGGTGTGCGACTCGAACCCG  
GCCACGGACTCCAAGGGCTCCTCTTCTCCCGCTGGGGATATCGGTCCGGGCGGCCAAC  
TCCAAGCTCGCCTTCTCGGCGGTGCGGAGCACCAACCACGAGCCATCCGAGATGAGCAAC  
AAGACGCGCATCATTTACTTCGATCAGATCCTGGTGAATGTGGGTAAATTTTTTCACATTG  
GAGTCTGTCTTTGTAGCACCAAGAAAAGGAATTTACAGTTTCAGTTTTTCACGTGATTA  
GTCTACCAGAGCCAAACATATCCAGGTTAACTTGATGTTAAATGGAAAACCAAGTAATCT  
GCCTTTGCGGGGGACAAAGATGTTACTCGTGAAGCTGCCACGAATGGTGTCTCTCTAC  
CTAGATAAAGAGGATAAGGTTTACCTAAAACCTGGAGAAAGGTAATTTGGTTGGAGGCTGG  
CAGTATTCACGTTTTCTGGCTTTCTGGTGTTCCTCCCTATAGGATTCAATTTCTCCATGA  
TGTTTCATCCAGGTGAGGGATGACCCACTCCTGAGTTATTGGAAGATCATTTTTTTCATCAT  
TGGATTGATGTCTTTTATGGTTTCTCATGGGTGGATATGGATTCTAAGGATTCTAGCCT  
GTCTGAACCAATACAAAATTTACAGATTATTTGTGTGTGTCTGTTTCAGTATATTTGGA  
TTGGGACTCTAAGCAGATAATACCTATGCTTAAATGTAACAGTCAAAGCTGTCTGCAAG  
ACTTATTTCTGAATTTCAATTCCTGGGATTACTGAATTAGTTACAGATGTGGAATTTTATT  
TGTTTAGTTTTTAAAAGACTGGCAACCAGGTCTAAGGATTAGAAAACCTCTAAAGTTCTGAC  
TTCAATCAACGGTTAGTGTGATACTGCCAAAGAACTGTATACTGTGTTAATATATTGATT  
ATATTTGTTTTTATTCCTTTGGAATTAGTTTGTGTTTGGTCTTGTAATAAACTTGGATTTT  
TTTTTTCAGTAACTGGTATTATGTTTTCTCTTAAATAAGGTAATGAATGGCTTGCCCCAC  
AAATTTACCTTGACTACGATATCATCGACATGACTTCTCTCAAAAAAAGAATGCTTCA  
TAGTTGTATTTTAATTGTATATGTGAAAGAGTCATATTTTCCAAGTTATATTTTCTAAGA  
AGAAGAATAGATCATAAATCTGACAAGGAAAAAGTTGCTTACCCAAAATCTAAGTGCTCA  
ATCCCTGAGCCTCAGCAAAACAGCTCCCCCTCCGAGGGAAATCTTATACTTTATTGCTCAA  
CTTTAATTTAAATGATTGATAATAACCACTTTATTAAAAACCTAAGGTTTTTTTTTTTTC  
CGTAGACATGACCACTTTATTAACGGTGGTGGGATGCTGTTGTTTCTAATTATACCTAT  
TTTTCAAGGCTTCTGTTGTATTTGAAGTATCATCTGGTTTTGCCTTAACTCTTTAAATTG  
TATATATTTATCTGTTTAGCTAATATTAAATTCAAATATCCCATATCTAAATTTAGTGCA  
ATATCTTGTCTTTTGTATAGGTCATATGAATTCATAAAATTATTTATGTCTGTTATAGAA  
TAAAGATTAATATATGTTAAAAAAA



225/392

## FIGURE 216

MSGRRALSAVPAVLLVLTLPGLPVWAQNDTEPIVLEGKCLVVCDSPATDSKGSSSSPL  
GISVRAANSKVAFSAVRSTNHEPSEMSNKTRIIYFDQILVNVGNFFTTLESV FVAPRKG IY  
SFSFHVIKVYQSQTIQVNLMLNGKPVISAFAGDKDVTREAA TNGVLLYLDKEDKVYLKLE  
KGNLVGGWQYSTFSGFLVFPL

Signal peptide:  
amino acids 1-27

226/392

**FIGURE 217**

CGGCAACCAGCCGCCGCCACCACCGCTGCCACTGCCGCCCTGCCGGGGCCATGTTTCGCTC  
TGGGCTTGCCCTTCTTGGTGCTCTTGGTGGCCTCGGTCGAGAGCCATCTGGGGGTTCTGG  
GGCCCAAGAACGTCTCGCAGAAAGACGCCGAGTTTGAGCGCACCTACGTGGACGAGGTCA  
ACAGCGAGCTGGTCAACATCTACACCTTCAACCATACTGTGACCCGCAACAGGACAGAGG  
GCGTGCGTGCTGTCTGTGAACGTCTTGAACAAGCAGAAGGGGGCGCCGTTGCTGTTTGTGG  
TCCGCCAGAAGGAGGCTGTGGTGTCCTTCCAGGTGCCCCAATCCTGCGAGGGATGTTTC  
AGCGCAAGTACCTCTACCAAAAGTGGAACGAACCCTGTGTGTCAGCCCCCACCAGAAGATG  
AGTCGGAGATTGAGTTCTTCTACGTGGATGTGTCCACCCTGTCAACAGTCAACACCACAT  
ACCAGCTCCGGGTGAGCCGCATGGACGATTTTGTGCTCAGGACTGGGGAGCAGTTCAGCT  
TCAATACCACAGCAGCACAGCCCCAGTACTTCAAGTATGAGTTCCTTGAAGGCGTGGACT  
CGGTAATTGTCAAGGTGACCTCCAACAAGGCCTTCCCCTGCTCAGTCATCTCCATTCAGG  
ATGTGCTGTGCTCTATGACCTGGACAACAACGTAGCCTTCATCGGCATGTACCAGA  
CGATGACCAAGAAGGCGGCCATCACCGTACAGCGCAAAGACTTCCCCAGCAACAGCTTTT  
ATGTGGTGGTGGTGGTGAAGACCGAAGACCAAGCCTGCCGGGGCTCCCTGCTTTCTACC  
CCTTCGCAGAAGATGAACCGGTGATCAAGGGCACCGCCAGAAAACCCCTGTGCTGCTGG  
TGTCTCAAGCAGTCACGTCTGAGGCATACGTGAGTGGGATGCTCTTTTGCCTGGGTATAT  
TTCTCTCCTTTTACCTGCTGACCGTCTCCTGGCCTGCTGGGAGAACTGGAGGCAGAAGA  
AGAAGACCCTGCTGGTGGCCATTGACCGAGCCTGCCAGAAAGCGGTCACCCTCGAGTCC  
TGGCTGATTCTTTTCTGGCAGTTCCCCTTATGAGGGTTACAACATATGGCTCCTTTGAGA  
ATGTTTCTGGATCTACCGATGGTCTGGTTGACAGCGCTGGCACTGGGGACCTCTCTTACG  
GTTACCAGGGCCGCTCCTTTGAACCTGTAGGTACTCGGCCCGAGTGGACTCCATGAGCT  
CTGTGGAGGAGGATGACTACGACACATTGACCGACATCGATTCCGACAAGAATGTCATTC  
GCACCAAGCAATACCTCTATGTGGCTGACCTGGCACGGAAGGACAAGCGTGTTCTGCGGA  
AAAAGTACCAGATCTACTTCTGGAACATTGCCACCATTGCTGTCTTCTATGCCCTTCCTG  
TGGTGCAGCTGGTGATCACCTACCAGACGGTGGTGAATGTCACAGGGAATCAGGACATCT  
GCTACTACAACCTTCTGCGCCCACTGGGCAATCTCAGCGCCTTCAACAACATCC  
TCAGCAACCTGGGGTACATCCTGCTGGGGCTGCTTTTCTGCTCATCATCTGCAACGGG  
AGATCAACCACAACCGGGCCCTGCTGCGCAATGACCTCTGTGCCCTGGAATGTGGGATCC  
CCAAACACTTTGGGCTTTTCTACGCCATGGGCACAGCCCTGATGATGGAGGGGCTGCTCA  
GTGCTTGCTATCATGTGTGCCCCAACTATAACCAATTTCCAGTTTGACACATCGTTCATGT  
ACATGATCGCCGACTCTGCATGCTGAAGCTCTACCAGAAGCGGCACCCGACATCAACG  
CCAGCGCCTACAGTGCCTACGCCTGCCTGGCCATTGTGTCATCTTCTTCTGTGCTGGGCG  
TGGTCTTTGGCAAAGGGAACACGGCGTTCTGGATCGTCTTCTCCATCATTCACATCATCG  
CCACCCTGCTCCTCAGCACGCAGCTCTATTACATGGGCGGGTGGAACTGGACTCGGGGA  
TCTTCCGCCGCATCCTCCACGTGCTCTACACAGACTGCATCCGGCAGTGCAGCGGGCCGC  
TCTACGTGGACCGCATGGTGCTGCTGGTCAATGGGCAACGTCATCAACTGGTCTGCTGGCTG  
CCTATGGGCTTATCATGCGCCCCAATGATTTGCTTCTTCTTCTTGTGGCCATTGGCATCT  
GCAACCTGCTCCTTTACTTCGCCTTCTACATCATCATGAAGCTCCGGAGTGGGGAGAGGA  
TCAAGCTCATCCCCCTGCTCTGCATCGTTTGACCTCCGTGGTCTGGGGCTTCGCGCTCT  
TCTTCTTCTTCCAGGGACTCAGCACCTGGCAGAAAACCCCTGCAGAGTGGAGGGAGACA  
ACCGGGACTGCATCCTCCTCGACTTCTTTGACGACCACGACATCTGGCACTTCTCTCCT  
CCATCGCCATGTTTCGGGTCTTCTGCTGTTGCTGACACTGGATGACGACCTGGATACTG  
TGCAGCGGGACAAGATCTATGTCTTCTAGCAGGAGCTGGGCCCTTCGCTTCACTCAAGG  
GGCCCTGAGCTCCTTTGTGTCATAGACCGGTCACTCTGTGCTGCTGTGGGGATGAGTCCC  
AGCACCGCTGCCAGCACTGGATGGCAGCAGGACAGCCAGGTCTAGCTTAGGCTTGGCCT  
GGGACAGCCATGGGGTGGCATGGAACCTTGCAGCTGCCCTCTGCCGAGGAGCAGGCCTGC  
TCCCCTGGAACCCCCAGATGTTGGCCAAATGCTGCTTTCTTCTCAGTGTGGGGCCTTC

227/392

CATGGGCCCCTGTCCTTTGGCTCTCCATTTGTCCCTTTGCAAGAGGAAGGATGGAAGGGA  
CACCCCTCCCCATTTTCATGCCTTGCATTTTGCCCGTCCTCCTCCCCACAATGCCCCAGCCT  
GGGACCTAAGGCCTCTTTTCTCCTCCATACTCCCACTCCAGGGCCTAGTCTGGGGCCTGA  
ATCTCTGTCCTGTATCAGGGCCCCAGTTCTCTTTGGGCTGTCCCTGGCTGCCATCACTGC  
CCATTCCAGTCAGCCAGGATGGATGGGGGTATGAGATTTTGGGGGTGGCCAGCTGGTGC  
CAGACTTTTGGTGCTAAGGCCTGCAAGGGGCCTGGGGCAGTGCGTATTCTCTTCCCTCTG  
ACCTGTGCTCAGGGCTGGCTCTTTAGCAATGCGCTCAGCCCAATTTGAGAACCGCCTTCT  
GATTCAAGAGGCTGAATTCAGAGGTCACCTCTTCATCCCATCAGCTCCCAGACTGATGCC  
AGCACCAGGACTGGAGGGAGAAGCGCCTCACCCCTTCCCTTCCTTCTTTCCAGGCCCTTA  
GTCTTGCCAAACCCAGCTGGTGGCCTTTCAGTGCCATTGACACTGCCCCAAGAATGTCCA  
GGGGCAAAGGAGGGATGATACAGAGTTCAGCCCGTTCTGCCTCCACAGCTGTGGGCACCC  
CAGTGCTTACCTTAGAAAGGGGCTTCAGGAAGGGATGTGCTGTTTCCCTCTACGTGCCCA  
GTCCTAGCCTCGCTCTAGGACCCAGGGCTGGCTTCTAAGTTTCCGTCCAGTCTTCAGGCA  
AGTTCTGTGTTAGTCATGCACACACATACCTATGAAACCTTGGAGTTTACAAAGAATTGC  
CCCAGCTCTGGGCACCCCTGGCCACCCCTGGTCCTTGGATCCCCCTTCGTCCCACCTGGTCCA  
CCCCAGATGCTGAGGATGGGGGAGCTCAGGCGGGGCCTCTGCTTTGGGGATGGGAATGTG  
TTTTTCTCCAACTTGTTTTTATAGCTCTGCTTGAAGGGCTGGGAGATGAGGTGGGTCT  
GGATCTTTTCTCAGAGCGTCTCCATGCTATGGTTGCATTTCCGTTTTCTATGAATGAATT  
TGCATTCAATAAACAACCAGACTCAAAAAAAAAAAAAA

228/392

**FIGURE 218**

MFALGLPFLVLLVASVESHLGVLGPKNVSQKDAEFERTYVDEVNSELVNIYTFNHTVTRN  
RTEGVRVSVNVLNKQKGAPLLFVVRQKEAVVSFQVPLILRGMFORKYLYQKVERTLCQPP  
TKNESEIQFFYVDVSTLSPVNTTYQLRVSRMDDFVLRTGEQFSFNTTAAQPQYFKYEFPE  
GVDSVIVKVTSNKAFFPCSVISIQDVLCPVYDLNNAFIGMYQTMKKAAITVQRKDFPS  
NSFYVVVVVKTEDQACGGSLPFYPFAEDEPVDQGHRQKTLVLVSQAVTSEAYVSGMLFC  
LGIFLSFYLLTVLLACWENWRQKKKTLLVAIDRACPESGHPRVLADSFPGSSPYEGYNYG  
SFENVSGSTDGLVDSAGTGDLSYGYQGRSFEPVGTTRPRVDSMSSVEEDDYDTLTDIDSDK  
NVIRTKQYLYVADLARKDKRVLRRKKYQIYFWNIATIAVFYALPVVQLVITYQTVVNVNVTGN  
QDICYYNFLCAHPLGNLSAFNNILSNLGYILLGLLFLLIILQREINHNRRALLRNDLCALE  
CGIPKHFGLFYAMGTALMMEGLLSACYHVCNNTNFQFDTSFMYMIAGLCMLKLYQKRHP  
DINASAYSAYACLAIVIFFSVLGVVFGKGNTAFWIVFSIIHIATLLLLSTQLYYMGRWKL  
DSGIFRRILHVLTYDCIRQCSGPLYVDRMVLLVMGNVINWSLAAYGLIMRPNDFASYLLA  
IGICNLLLYFAFYIIMKLRSGERIKLIPLLCIVCTSVVWGFALFFFFQGLSTWQKTPAES  
REHNRDCILLDFDDHDIWHFLSSIAMFGSFLVLLTLDDDLDTVQRDKIYVF

**Important features of the protein:****Signal peptide:**

amino acids 1-18

**Transmembrane domains:**

amino acids 292-317, 451-470, 501-520, 607-627, 751-770

**Leucine zipper pattern:**

amino acids 497-518

**N-glycosylation sites:**amino acids 27-30, 54-57, 60-63, 123-126, 141-144, 165-168,  
364-367, 476-479, 496-499, 572-575, 603-606, 699-702



230/392

## FIGURE 220

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQML  
TLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTLISSEELP  
QIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSGTDD  
DFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:  
amino acids 1-16

231/392

**FIGURE 221**

GACTTTGCTTGAATGGTTTACATATTTCTGCTCGCTGTCTCATATCACAAATATAGTCTT  
ACGTTTTTGTAAAACTTTGGGGTGTGAGGAGTTGAGCTTGCTCAGCAAGCCAGCATGGCT  
AGGATGAGCTTTGTTATAGCAGCTTGCCAATTGGTGCTGGGCCTACTAATGACTTTCATTA  
ACCGAGTCTTCCATACAGAATAGTGAGTGCCACAACCTTTGCGTATGTGAAATTCGTCCC  
TGGTTTACCCACAGTCAACTTACAGAGAAGCCACCACTGTTGATTGCAATGACCTCCGC  
TTAACAAGGATTCCCAGTAACCTCTCTAGTGACACACAAGTGCTTCTCTTACAGAGCAAT  
AACATCGCGAAGACTGTGGATGAGCTGCAGCAGCTTTTCAACTTGACTGAACTAGATTTT  
TCCAAAACAACCTTTACTAACATTAAGGAGGTGCGGCTGGCAAACCTAACCCAGCTCAC  
ACGCTGCATTTGGAGGAAAATCAGATTACCGAGATGACTGATTACTGTCTACAAGACCTC  
AGCAACCTTCAAGAACTCTACATCAACCACAACCAAATTAGCACTATTTCTGCTCATGCT  
TTTGACAGGCTTAAAAAATCTATTAAGGCTCCACCTGAACTCCAACAAATTGAAAGTTATT  
GATAGTCGCTGGTTTGATTCTACACCCAACCTGGAAATTCTCATGATCGGAGAAAACCT  
GTGATTGGAATTCTGGATATGAACTTCAAACCCCTCGCAAATTTGAGAAGCTTAGTTTTG  
GCAGGAATGTATCTCACTGATATTCCTGGAAATGCTTTGGTGGGTCTGGATAGCCTTGAG  
AGCCTGTCTTTTTATGATAACAACTGGTTAAAGTCCCTCAACTTGCCCTGCAAAAAGTT  
CCAAATTTGAAATTTCTAGACCTCAACAAAAACCCATTCAAAAAATCCAAGAAGGGGAC  
TTCAAAAATATGCTTCGGTTAAAAGAACTGGGAATCAACAATATGGGCGAGCTCGTTTTCT  
GTCGACCGCTATGCCCTGGATAACTTGCCTGAACTCACAAAGCTGGAAGCCACCAATAAC  
CCTAACTCTCTTACATCCACCGCTTGGCTTTCGGAAGTGTCCCTGCTCTGGAAAGCTTG  
ATGCTGAACAACAATGCCCTGAATGCCATTTACCAAAAGACAGTCGAATCCCTCCCAAT  
CTGCGTGAGATCAGTATCCATAGCAATCCCTCAGGTGTGACTGTGTGATCCACTGGATT  
AACTCCAACAAAACCAACATCCGCTTATGGAGCCCTGTCCATGTCTGTGCCATGCGG  
CCCGAATATAAAGGGCACCAGGTGAAGGAAGTTTAAATCCAGGATTCGAGTGAACAGTGC  
CTCCCAATGATATCTCACGACAGCTTCCCAAATCGTTTTAAACGTGGATATCGGCACGACG  
GTTTTCTTAGACTGTGAGCCATGGCTGAGCCAGAACCTGAAATTTACTGGGTCACTCCC  
ATTGGAATAAGATAACTGTGGAAACCTTTTCAGATAAATACAAGCTAAGTAGCGAAGGT  
ACCTTGGAATATCTAACATACAAATTGAAGACTCAGGAAGATACACATGTGTTGCCAG  
AATGTCCAAGGGGCAGACACTCGGGTGGCAACAATTAAAGGTTAACGGGACCTTCTGGAT  
GGTACCCAGGTGCTAAAAATATACGTCAAGCAGACAGAATCCCATTCATCTTAGTGTCC  
TGGAAAGTTAATTCCAATGTGATGACGTCAAACCTTAAATGGTGTCTGCCACCATGAAG  
ATTGATAACCCCTCACATAACATATACTGCCAGGGTCCAGTCGATGTCCATGAATACAAC  
CTAACGCATCTGCAGCCTTCCACAGATTATGAAGTGTGTCTCACAGTGTCCAATATTCAT  
CAGCAGACTCAAAAGTCATGCGTAAATGTACAACCAAAAATGCCGCTTCGCAGTGGAC  
ATCTCTGATCAAGAAACCAGTACAGCCCTTGCTGCAGTAATGGGGTCTATGTTTGCCGTC  
ATTAGCCTTGCGTCCATTGCTGTGTAATTTGCCAAAAGATTTAAGAGAAAAAACTACCAC  
CACTCATTAACCTCTGGGAAGGTGACAGCGAGAAAGACAAAGATGGTTCTGCAGACACC  
AAGCCAACCCAGGTGACACATCCAGAAGCTATTACATGTGGTAACTCAGAGGATATTTT  
GCTTCTGGTAGTAAGGAGCACAAAGACGTTTTTGTCTTTATTTCTGCAAAAGTGAACAAGTT  
GAAGACTTTTGTATTTTTGACTTTTGCTAGTTTGTGGCAGAGTGGAGAGGACGGGTGGATA  
TTTTCAAAATTTTTTAGTATAGCGTATCGCAAGGGTTTGACACGGCTGCCAGCGACTCTAG  
GCTTCCAGTCTGTGTTTTGGTTTTTATTCTTATCATTATTATGATTGTTATTATATTATTA  
TTTTATTTTGTGTTGTTGTGCTAAACTCAATAATGCTGTTCTAACTACAGTGCTCAATAAA  
ATGATTAATGACAGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG  
AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG

232/392

**FIGURE 222**

MARMSFVIAACQLVLGLLMTSLTESSIONSECPQLCVCEIRPWFTPQSTYREATTVDCND  
LRLTRIPSNLSSDTQVLLQLSQNNIAKTVDLQQLFNLTELDQSQNNFTNIKEVGLANLTQ  
LTTLHLEENQITEMTDYCLQDLNLQELYINHNQISTISAHAFAGLKNLLRLHLNSNKLK  
VIDSRWFDSTPNLEILMIGENPVIGILDMNFKPLANLRSLVLAGMYLTDIPGNALVGLDS  
LESLSFYDNKLVKVPQLALQKVPNLKFLDLNKNPIHKIQEGDFKNMLRLKELGINNMGEL  
VSVDRYALDNLPELTKLEATNNPKLSYIHRLAFRSVPALESMLNNAALNAIYQKTVESL  
PNLREISIHNSNPLRCDLCVIHWINSNKTNIRFMEPLSMFCAMPPEYKGHVKEVLIQDSSE  
QCLPMISHDSFPNRLNVDIGTTVFLDCRAMAEPEPEIYWVTPIGNKITVETLSDKYKLSS  
EGTLEISNIQIEDSGRYTCVAQNVQGADTRVATIKVNGTLLDGTQVLKIYVKQTESHSIL  
VSWKVSNSVMTSNLKWSSATMKIDNPHITYTARVPVDVHEYNLTHLQPSTDYEVCLTVSN  
IHQQTQKSCVNVTTKNAFAVDISDQETSTALAAVMGSMFAVISLASIAVYFAKRFRKRN  
YHSLKMYMQKTSSIPLNELYPPLINLWEGDSEKDKDGSADTKPTQVDTSRSYMW

**Important features:****Signal peptide:**

Amino acids 1-25

**Transmembrane domain:**

Amino acids 508-530

**N-glycosylation sites:**Amino acids 69-73;96-100;106-110;117-121;385-389;517-521;  
582-586;611-615**Tyrosine kinase phosphorylation site:**

Amino acids 573-582

**N-myristoylation sites:**

Amino acids 16-22;224-230;464-470;637-643;698-704



233/392

**FIGURE 223**

[illegible]

234/392

**FIGURE 224**

MASYLYGVLFVAVGLCAPIYCVSPANAPSAYPRPSSTKSTPASQVYSLNTDFAFRLYRRLV  
LETPSQNIFFSPVSVSTSLAMLSLGAHSVTKTQILQGLGFNLTHTPESAIHQGFQHLVHS  
LTVPSKDLTLKMGSALFVKKELQLQANFLGNVKRLYEAEVFSTDFSNPSIAQARINSHVK  
KKTQGKVVDIIQGLDLLTAMVLVNHIFKAKWEKPFHLEYTRKNFPFLVGEQVTVQVPMM  
HQKEQFAFGVDTELNCFVLQMDYKGDVAFFVLPSKKGMRQLEQALSARTLIKWSHSLQK  
RWIEVFIPRFSISASYNLETILPKMGIQNAFDKNADFSGIKRDSLQVSKATHKAVLDVS  
EEGTEATAATTTKFIVRSKDGPSYFTVSFNRTFLMMITNKATDGILFLGKVENPTKS

**Signal peptide:**  
amino acids 1-20

235/392

**FIGURE 225**

GGGAGAGAGGATAAATAGCAGCGTGGCTTCCCTGGCTCCTCTCTGCATCCTTCCCGACCT  
TCCCAGCAATATGCATCTTGACAGTCTGGTCGGCTCCTGCTCCCTCCTTCTGCTACTGGG  
GGCCCTGTCTGGATGGGCGGCCAGCGATGACCCCATTGAGAAGGTCATTGAAGGGATCAA  
CCGAGGGCTGAGCAATGCAGAGAGAGAGGTGGGCAAGGCCCTGGATGGCATCAACAGTGG  
AATCACGCATGCCGGAAGGGAAGTGGAGAAGGTTTTCAACGGACTTAGCAACATGGGGAG  
CCACACCGGCAAGGAGTTGGACAAAGGCGTCCAGGGGCTCAACCACGGCATGGACAAGGT  
TGCCCATGAGATCAACCATGGTATTGGACAAGCAGGAAAGGAAGCAGAGAAGCTTGGCCA  
TGGGGTCAACAACGCTGCTGGACAGGCCGGAAGGAAGCAGACAAAGCGGTCCAAGGGTT  
CCCACTGGGGTCCACCAGGCTGGGAAGGAAGCAGAGAACTTGGCCAAGGGGTCAACCA  
TGCTGCTGACCAGGCTGGAAAGGAAGTGGAGAAGCTTGGCCAAGGTGCCACCATGCTGC  
TGGCCAGGCCGGAAGGAGCTGCAGAATGCTCATAATGGGGTCAACCAAGCCAGCAAGGA  
GGCCAACCAGCTGCTGAATGGCAACCATCAAAGCGGATCTTCCAGCCATCAAGGAGGGGC  
CACAACCACGCCGTTAGCCTCTGGGGCCTCAGTCAACACGCCTTTCATCAACCTTCCCGC  
CCTGTGGAGGAGCGTCGCCAACATCATGCCCTAAACTGGCATCCGGCCTTGCTGGGAGAA  
TAATGTCGCCGTTGTACATCAGCTGACATGACCTGGAGGGGTGGGGGTGGGGGACAGG  
TTTCTGAAATCCCTGAAGGGGGTTGTAAGGGATTTGTGAATAAACTTGATACACCA

236/392

## FIGURE 226

MHLARLVGSCSLLLLLLGALSGWAASDDPIEKVIEGINRGLSNAEREVKGALDGINSGITH  
AGREVEKVFNGLSNMGSHGTGKELDKGVQGLNHGMDKVAHEINHGIGQAGKEAEKLGHG  
VN  
NAAGQAGKEADKAVQGFHTGVHQAGKEAEKLGQGVNHAADQAGKEVEKLGQGAHHAAGQA  
GKELQNAHNGVNVQASKEANQLLNGNHQSGSSSHQGGATTTPLASGASVNTPFINLPALWR  
SVANIMP

**Important features of the protein:**

**Signal peptide:**

amino acids 1-25

**Homologous region to circumsporozoite (CS) repeats:**

amino acids 35-225

237/392

**FIGURE 227**

GAAGTAGAGGTGTTGTGCTGAGCGGCGCTCGGCGAACTGTGTGGACCGTCTGCTGGGACT  
CCGGCCCTGCGTCCGCTCAGCCCCGTGGCCCCGCGCACCTACTGCCATGGGAGACGCGGCC  
TCGTCTCGGGGCCACCTGTTTGCTGGGCTTCAGTTTCCTGCTCCTCGTCATCTCTTCTGA  
TGGACATAATGGGCTTGGAAGGGTTTTGGAGATCATATTCATTGGAGGACACTGGAAGA  
TGGGAAGAAAGAAGCAGCTGCCAGTGGACTGCCCCGTGATGGTGATTATTCATAAATCCTG  
GTGTGGAGCTTGCAAAGCTCTAAAGCCCCAAATTTGCAGAATCTACGGAAATTTCAGAACT  
CTCCCATAAATTTTGTTATGGTAAATCTTGAGGATGAAGAGGAACCCAAAGATGAAGATTT  
CAGCCCTGACGGGGGTATATTCACGAATCCTTTTTCTGGATCCCAGTGGCAAGGTGCA  
TCCTGAAATCATCAATGAGAATGGAAACCCAGCTACAAGTATTTTTATGTCAGTGCCGA  
GCAAGTTGTTCAAGGGATGAAGGAAGCTCAGGAAAGGCTGACGGGTGATGCCTTCAGAAA  
GAAACATCTTGAAGATGAATTGTAACATGAATGTGCCCCCTTCTTTCATCAGAGTTAGTGT  
TCTGGAAGGAAAGCAGCAGGGAAGGGAATATTGAGGAATCATCTAGAACAATTAAGCCGA  
CCAGGAAACCTCATTCCTACCTACACTGGAAGGAGCGCTCTCACTGTGGAAGAGTTCTGC  
TAACAGAAGCTGGTCTGCATGTTTGTGGATCCAGCGGAGAGTGGCAGACTTTCTTCTCCT  
TTTCCCTCTCACCTAAATGTCAACTTGTCATTGAATGTAAAGAATGAAACCTTCTGACAC  
AAAA

238/392

## FIGURE 228

METRPRLGATCLLGFSFLLLVISSDGHNGLGKGFGDHIHWRTLEDGKKEAAASGLPLMVI  
IHKSWCGACKALKPKFAESTEISELSHNFVMVNLEDEEPPKDEDFSPDGGYIPRILFLDP  
SGKVHPEIINENGNP SYKYFYVSAEQVVQGMKEAQERLTGDAFRKKHLEDEL

Signal peptide:  
Amino acids 1-23

Thioredoxin family proteins Homology Block:  
Amino acids 58-75

239/392

**FIGURE 229**

CCCACGCGTCCGCCCACGCGTCCGGGTGCCACTCGCGCGCCGGCCGCGCTCCGGGCTTCT  
CTTTTCCCTCCGACGCGCCACGGCTGCCCAGACATTCCGGCTGCCGGGTCTGGAGAGCTC  
CCCGAACCCCTCCGCGGAGAGGAGCGAGGCGGCGCCAGGGTGGCCCCCGGGGCGCGCTTG  
GTCTCGGAGAAGCGGGGACGAGGCCGGAGGATGAGCGACTGAGGGCGACGCGGGCACTGA  
CGCGAGTTGGGGCCGCGACTACCGGCAGCTGACAGCGCGATGAGCGACTCCCCAGAGACG  
CCCTAGCCCCGGTGTGCGCGCCAGGCGGAGCGCGCAGGTGGGGCTGGGCTGTTAGTGGTCC  
GCCCCACGCGGGTCGCCGGCCGGCCAGGATGGGCGCTGGCAACCCGGGCCCCGCGCCCGC  
CGCTGCTACCCCTGCGCCCGCTGCGAGCCCGGCGTCCGGCCCCGCGCCCTGCGCTCATGGA  
CGGCGGCTCCCGGCTGGCGGCGGCGCGCCCCCGGGCTGTGAATGCGACTCGCCCCCTCGGC  
CGCGCTCCCCGCCCCGCCCCGCGGGGAGTGGTAGGGGATGCCAGCTCCACTGCGAT  
GGCAGTTGGCGCGCTCTCCAGTTCCCTCCTGGTCACTGCTGCCTGATGGTGGCTCTGTG  
CAGTCCGAGCATCCCGCTGGAGAAGCTGGCCAGGCACCGAGCAGCGGGCCAGAGAGAA  
GCGTGAGCACGCCACTCGGGACGGCCCCGGGGCGGGTGAACGAGCTCGGGCGCCCGGCGCG  
GGACGAGGGCGGCAGCGGCCGGGACTGGAAGAGCAAGAGCGGCCGTGGGCTCGCCGGCCG  
TGAGCCGTGGAGCAAGCTGAAGCAGGCCCTGGGTCTCCAGGGCGGGGGCGCCAAGGCCGG  
GGATCTGCAGGTCCGGCCCCGCGGGACACCCCGCAGGCGGAAGCCCTGGCCGCAGCCGC  
CCAGGACGCGATTGGCCCCGGAATCGCGCCCACGCCCCGAGCCACCCGAGGAGTACGTGTA  
CCCGGACTACCGTGGAAGGGCTGCGTGACGAGAGCGGCTTCGTGTACGCGATCGGGGA  
GAAGTTCGCGCCGGGCCCCCTCGGCCCTGCCGTGCCTGTGCACCGAGGAGGGCCGCTGTG  
CGCGCAGCCCCGAGTGCCCGAGGCTGCACCCGCGCTGCATCCACGTGACACGAGCCAGTG  
CTGCCCCGAGTGCAAGGAGAGGAAGAATACTGCGAGTTCCGGGGCAAGACCTATCAGAC  
TTTGAGGAGTTCGTGGTGTCTCCATGCGAGAGGTGTCGCTGTGAAGCCAACGGTGAGGT  
GCTATGCACAGTGTGAGCGTGTCCCCAGACGGAGTGTGTGGACCCTGTGTACGAGCCTGA  
TCAGTGCTGTCCCATCTGCAAAAATGGTCCAACTGCTTTGCAGAAACCGCGGTGATCCC  
TGCTGGCAGAGAAGTGAAGACTGACGAGTGCACCATATGCCACTGTACTTATGAGGAAGG  
CACATGGAGAATCGAGCGGCAGGCCATGTGCACGAGACATGAATGCAGGCAAAATGTAGAC  
GCTTCCCAGAACACAACTCTGACTTTTTCTAGAACATTTTACTGATGTGAACATTCTAG  
ATGACTCTGGGAACTATCAGTCAAAGAAGACTTTTGATGAGGAATAATGGAAAATTGTTG  
GTACTTTTCTTTTCTTGATAACAGTTACTACAACAGAAGGAAATGGATATATTTCAAAA  
CATCAACAAGAACTTTGGGCATAAAATCCTTCTCTAAATAAATGTGCTATTTTCACAGTA  
AGTACACAAAAGTACACTATTATATATCAAATGTATTTCTATAATCCCTCCATTAGAGAG  
CTTATATAAGTGTTTTCTATAGATGCAGATTAAAAATGCTGTGTTGTCAACCGTCAAAAA  
AAAAAAAAAAAAAAAAAAAAA

240/392

**FIGURE 230**

MPSSTAMAVGALSSSLLVTCCLMVALCSPSIPLEKLAQAPEQPGQEKREHATRDGPGRVN  
ELGRPARDEGGSGRDWKSXSGRGLAGREPWSKLKQAWVSQGGGAKAGDLQVRPRGDTPOA  
EALAAAAQDAIGPELAPTPEPPPEEYVYPDYRGKGCVDSESGFVYAIGKFAFGPSACPCLC  
TEEGPLCAQPECPRLHPRCIHVDTSQCCPQCKERKNYCEFRGKTYQTLEEFVVSPCERCR  
CEANGEVLCTVSACPQTECVDPVYEPDQCCPICKNGPNCFAETAVIPAGREVKTDECTIC  
HCTYEETWRIERQAMCTRHECRQM

**Important features of the protein:**

**Signal peptide:**

amino acids 1-27

**Transmembrane domain:**

amino acids 11-30

**Glycosaminoglycan attachment site:**

amino acids 80-83

**N-myristoylation sites:**

amino acids 10-15, 102-107, 103-108

**Cell attachment sequence:**

amino acids 114-117

**EGF-like domain cysteine pattern signature:**

amino acids 176-187



241/392

**FIGURE 231**

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGAC  
CCCTACGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTT  
GTCACAGGTGGGAGGCTGGAACTATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGG  
AAGCTGGCAACAAATGGATGATGTGATATATGCAATTCCAGGGGAAGGGAAATTGTGGTG  
TTCTGAACCCATGGTCAATTAACGAGGCAGTTTTCTAGCTACTGCACGTACTTCATAAAGC  
AGGACTCTAAAAGCTTTTGAATCATGGTGTGATGGAAAGGGATTTACTTTTATACTGACTC  
TGTTTTGGGGAAGCTTTTTTGAAGCATTTCATGCTGAGTCCCTTTTTTACCTTTGATGT  
TTGTAAACCCATCTTGGTATCGCTGGATCAACAACCGCCTTGTGGCAACATGGCTCACCC  
TACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAAAGTGATTATAACTGGGGATGCAT  
TTGTTCTCGGAGAAAGAAGTGTCATTATCATGAACCATCGGACAAGAATGGACTGGATGT  
TCCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATTTGCCTCAAAG  
CGAGTCTCAAAGGTGTTTCTGGATTGGTTGGGCCATGCAGGCTGCTGCCTATATCTTCA  
TTCATAGGAAATGGAAGGATGACAAGAGCCATTTCGAAGACATGATTGATTACTTTTGTG  
ATATTCACGAACCACTTCAACTCCTCATATTTCCAGAAGGGACTGATCTCACAGAAAACA  
GCAAGTCTCGAAGTAATGCATTTGCTGAAAAAAATGGACTTCAGAAATATGAATATGTTT  
TACATCCAAGAACTACAGGCTTTACTTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACC  
TTGATGCTGTCCATGATATCACTGTGGCGTATCCTCACAACATTTCCTCAATCAGAGAAGC  
ACCTCCTCCAAGGAGACTTTCCAGGGAAATCCACTTTCACGTCCACCGGTATCCAATAG  
ACACCCTCCCCACATCCAAGGAGGACCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGA  
AAGAAGAGAGGCTGCGTTCTTCTATCAAGGGGAGAAGAATTTTTATTTTACCGGACAGA  
GTGTCATTCCACCTTGCAAGTCTGAACTCAGGGTCTTGTGGTCAAATTGCTCTCTATAC  
TGTATTGGACCTGTTTCAAGCTGCAATGTGCCTACTCATATATTTGTACAGTCTTGTTA  
AGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGAGAATATTTGGTGGAC  
TGGAGATCATAGAACTTGATGTTACCGACTTTTACACAAACAGCCACATTTAAATTCAA  
AGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTGGAAAT  
GTTCTAAACCTTTCTAAGCTCAGATGCATTTTTGTCATGACTATGTCGAATATTTCTTACT  
GCCATCATTATTTGTTAAAGATATTTTGCACCTAATTTTGTGGGAAAAATATTGCTACAA  
TTTTTTTTTAATCTCTGAATGTAATTTTCGATACTGTGTACATAGCAGGGAGTGATCGGGT  
GAAATAACTTGGGCCAGAATATTATTAAACAATCATCAGGCTTTTAAA

242/392

**FIGURE 232**

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSF GIMVSWKGIYFILTLFWGSFFGSI  
FMLSPFLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVII  
MNHRTMRMDWMFLWNCLMRYSYLRLEKICLKASLKGVPFGFGWAMQAAAYIFIHRKWKDDKS  
HFEDMIDYFCDIHEPLQLLIFPEGTDLTENSKSRSNFAEKNGLQKYEYVLHPRTTGFTF  
VVDRLREGKNLDAVHDITVAYPHNIPQSEKHLLOGDFPREIH FHVHRYPIDTLPTS KEDL  
QLWCHKRWEEKERLRSFYQGEKNFYFTGQSVIPPCKSELRLVVKLLSILYWTLFSPAM  
CLLIYLYSLVKWYFIITIVIFVLQERIFGGLEIIELACYRLLHKQPHLNSKKNE

**Important features of the protein:**

**Signal peptide:**

amino acids 1-22

**Transmembrane domains:**

amino acids 44-63, 90-108, 354-377

243/392

**FIGURE 233**

CGGCTCGAGTGCAGCTGTGGGGAGATTTCACTGCATTGCCTCCCCCTGGGTGCTCTTCATC  
TTGGATTTGAAAAGTTGAGAGCAGCATGTTTTGCCCCACTGAAACTCATCCTGCTGCCAGTG  
TTACTGGATTATTCCTTGGGCCTGAATGACTTGAATGTTTTCCCCGCCTGAGCTAACAGTC  
CATGTGGGTGATTTCAGCTCTGATGGGATGTGTTTTCCAGAGCACAGAAGACAAATGTATA  
TTCAAGATAGACTGGACTCTGTCACCAGGAGAGCACGCCAAGGACGAATATGTGCTATAC  
TATTACTCCAATCTCAGTGTGCCTATTGGGCGCTTCCAGAACC GCGTACACTTGATGGGG  
GACATCTTATGCAATGATGGCTCTCTCTGCTCCAAGATGTGCAAGAGGCTGACCAGGGA  
ACCTATATCTGTGAAATCCGCCTCAAAGGGGAGAGCCAGGTGTTCAAGAAGGCGGTGGTA  
CTGCATGTGCTTCCAGAGGAGCCCAAAGAGCTCATGGTCCATGTGGGTGGATTGATTCAG  
ATGGGATGTGTTTTCCAGAGCACAGAAGTGAAACACGTGACCAAGGTAGAATGGATATTT  
TCAGGACGGCGCGCAAAGGAGGAGATTGTATTTCGTTACTACCACAACTCAGGATGTCT  
GTGGAGTACTCCCAGAGCTGGGGCCACTTCCAGAATCGTGTGAACCTGGTGGGGGACATT  
TTCCGCAATGACGGTTCATCATGCTTCAAGGAGTGAGGGAGTCAGATGGAGGAACTAC  
ACCTGCAGTATCCACCTAGGGAACCTGGTGTTCAGAAAACCATTTGTGCTGCATGTGAGC  
CCGGAAGAGCCTCGAACACTGGTGACCCCGGCAGCCCTGAGGCCTCTGGTCTTGGGTGGT  
AATCAGTTGGTGATCATTGTGGGAATTGTCTGTGCCACAATCCTGCTGCTCCCTGTTCTG  
ATATTGATCGTGAAGAAGACCTGTGGAAATAAGAGTTCAGTGAATTCTACAGTCTTGGTG  
AAGAACACGAAGAAGACTAATCCAGAGATAAAAGAAAAACCTGCCATTTTGAAAGATGT  
GAAGGGGAGAAACACATTTACTCCCCAATAATTGTACGGGAGGTGATCGAGGAAGAAGAA  
CCAAGTGAAAAATCAGAGGCCACCTACATGACCATGCACCCAGTTTGGCCTTCTCTGAGG  
TCAGATCGGAACAACTCACTTGAAAAAAAGTCAGGTGGGGGAATGCCAAAAACACAGCAA  
GCCTTTTGAGAAGAATGGAGAGTCCCTTCATCTCAGCAGCGGTGGAGACTCTCTCCTGTG  
TGTGTCCTGGGCCACTCTACCAGTGATTTTCAGACTCCCGCTCTCCAGCTGTCTCCTGT  
CTCATTTGTTTTGGTCAATACACTGAAGATGGAGAATTTGGAGCCTGGCAGAGAGACTGGAC  
AGCTCTGGAGGAACAGGCCTGCTGAGGGGAGGGGAGCATGGACTTGGCCTCTGGAGTGGG  
ACACTGGCCCTGGGAACCAGGCTGAGCTGAGTGGCCTCAAACCCCCCGTTGGATCAGACC  
CTCCTGTGGGCAGGGTTCTTAGTGGATGAGTTACTGGGAAGAATCAGAGATAAAACCAA  
CCCAAATCAA

244/392

**FIGURE 234**

MFCPLKLILLPVLLDYSGLNDLNVSPPELTVHVGDSALMGCVFQSTEDKCIFKIDWTLS  
PGEHAKDEYVLYYYSNLSVPIGRFQNRVHLMGDILCNDGSLLLQDVQEADQGTYICEIRL  
KGESQVFKKAVVLHVLPEEPKELMVHVGGLIQMGCVFQSTEVKHVTKVEWIFSGRRAKEE  
IVFRYYHKL RMSVEYSQSWGHFQNRVNLVGDI FRNDGSIMLQGVRES DGGNYTCSIHLGN  
LVFKKTIVLHVSPEEPRTLVT PAALRPLVLGGNQLV IIVGIVCATILLLPVLILIVKKTC  
GNKSSVNSTVLVKNTKKTNP EIKEPCHFERCEGEKHIYSPIIVREVIEEEEPSEKSEAT  
YMTMHPVWPSLRSDRNN SLEKKSGGMPKTQQAF

245/392

**FIGURE 235**

TAAACAGCTACAATATTCAGGGCCAGTCACTTGCCATTTCTCATAACAGCGTCAGAGA  
GAAAGAACTGACTGAAACGTTTGAGATGAAGAAAGTTCTCCTCCTGATCACAGCCATCTT  
GGCAGTGGCTGTTGGTTTCCCACTCTCTCAAGACCAGGAACGAGAAAAAGAAGTATCAG  
TGACAGCGATGAATTAGCTTCAGGGTTTTTTGTGTTCCCTTACCCATATCCATTTGCCC  
ACTTCCACCAATTCCATTTCCAAGATTTCCATGGTTTAGACGTAATTTTCTATTCCAAT  
ACCTGAATCTGCCCCCTACAACTCCCCTTCCTAGCGAAAAGTAAACAAGAAGGATAAGTCA  
CGATAAACCTGGTCACCTGAAATTGAAATTGAGCCACTTCCTTGAAGAATCAAAATTCCT  
GTTAATAAAAGAAAAACAAATGTAATTGAAATAGCACACAGCATTCTCTAGTCAATATCT  
TTAGTGATCTTCTTTAATAAACATGAAAGCAAAGATTTTGGTTTCTTAATTTCCACA

246/392

## FIGURE 236

MKKVLLLLITAILAVAVGFPVSQDQEREKRSISDSDELASGFFVFPYPYPFRPLPPIFFPR  
FPWFRRNFPIPIPIESAPTTPLPSEK

Important features of the protein:

Signal peptide:

amino acids 1-17

Homologous region to B3-hordein:

amino acids 47-85

247/392

**FIGURE 237**

TCGCCATGGCCTCTGCCGGAATGCAGATCCTGGGAGTCGTCCTGACACTGCTGGGCTGGG  
TGAATGGCCTGGTCTCCTGTGCCCTGCCCATGTGGAAGGTGACCGCTTTCATCGGCAACA  
GCATCGTGGTGGCCCAGGTGGTGTGGGAGGGCCTGTGGATGTCCTGCGTGGTGCAGAGCA  
CCGGCCAGATGCAGTGCAAGGTGTACGACTCACTGCTGGCGCTGCCACAGGACCTGCAGG  
CTGCACGTGCCCTCTGTGTCAATCGCCCTCCTTGTGGCCCTGTTGGCTTGGTCTTACC  
TTGCTGGGGCCAAGTGTACCACTGTGTGGAGGAGAAGGATTCCAAGGCCCGCCTGGTGC  
TCACCTCTGGGATTGTCTTTGTCACTCAGGGGTCTGACGCTAATCCCCGTGTGCTGGA  
CGGCGCATGCCATCATCCGGGACTTCTATAACCCCTGGTGGCTGAGGCCCAAAAGCGGG  
AGCTGGGGGCCCTCCCTCTACTTGGGCTGGGCGGCCTCAGGCCTTTTGTGTGCTGGGTGGGG  
GGTTGCTGTGCTGCACTTGCCCCCTCGGGGGGTCCCAGGGCCCCAGCCATTACATGGCCC  
GCTACTCAACATCTGCCCCCTGCCATCTCTCGGGGGCCCTCTGAGTACCCTACCAAGAATT  
ACGTCTGACGTGGAGGGGAATGGGGGCTCCGCTGGCGCTAGAGCCATCCAGAAGTGGCAG  
TGCCCAACAGCTTTGGGATGGGTTCGTACCTTTTGTCTGCTCCTGCTATTTTTCTTT  
TGACTGAGGATATTTAAATTCATTTGAAAACAGGCAAGGTGTTGACTCAGACTCTCA  
CTTAGGCTCTGCTGTTTCTCACCCCTTGGATGATGGAGCCAAAGAGGGGATGCTTTGAGAT  
TCTGGATCTTGACATGCCATCTTAGAAGCCAGTCAAGCTATGGAACATAATGCGGAGGCT  
GCTTGCTGTGCTGGCTTTGCAACAAGACAGACTGTCCCCAAGAGTTCTGCTGCTGCTGG  
GGGCTGGGCTTCCCTAGATGTCACTGGACAGCTGCCCCCATCCTACTCAGGTCTCTGGA  
GCTCCTCTCTTACCCCTGGAAAAACAAATCATCTGTTAACAAGGACTGCCACCTCCG  
GAACTTCTGACCTCTGTTTCTCCGTCCTGATAAGACGTCCACCCCCAGGGCCAGGTCC  
CAGCTATGTAGACCCCCGCCCCACCTCCAACACTGCACCCTTCTGCCCTGCCCCCTCG  
TCTCACCCCTTTACACTCACATTTTTATCAAATAAAGCATGTTTTGTTAGTGCA

248/392

## FIGURE 238

MASAGMQILGVVLTLLGWVNGLVSCALPMWKVTAFIGNSIVVAQVVWEGWLWMSCVVQSTG  
QMCKVYDSLALLPQDLQAARALCVIALLVALLFGLLVYLAGAKCTTCVEEKDSKARLVLT  
SGIVFVISGVLTLIPVCWTAHAIIRDFYNPLVAEAQKRELGASLYLGWAASGLLLLGGGL  
LCCTCPSGGSQGPHSHYMARYSTSAPAIISRGPSYPTKNYV

**Transmembrane domains:**

amino acids 8-30 (type II), 82-102, 121-140, 166-186



249/392

**FIGURE 239**

AGTGACAATCTCAGAGCAGCTTCTACACCACAGCCATTTCCAGCATGAAAGATCACTGGGG  
GTCTCCTTCTGCTCTGTACAGTGGTCTATTTCTGTAGCAGCTCAGAAGCTGCTAGTCTGT  
CTCCAAAAAAGTGGACTGCAGCATTTACAAGAAGTATCCAGTGGTGGCCATCCCCTGCC  
CCATCACATACCTACCAGTTTGTGGTTCTGACTACATCACCTATGGGAATGAATGTCACT  
TGTGTACCGAGAGCTTGAAAAGTAATGGAAGAGTTCAGTTTCTTCACGATGGAAGTTGCT  
AAATTCTCCATGGACATAGAGAGAAAAGGAATGATATTCTCATCATCATCTTCATCATCCC  
AGGCTCTGACTGAGTTTCTTTCAGTTTTACTGATGTTCTGGGTGGGGGACAGAGCCAGAT  
TCAGAGTAATCTTGACTGAATGGAGAAAGTTTCTGTGCTACCCCTACAAACCCATGCCTC  
ACTGACAGACCAGCATTTTTTTTTTAACACGTCAATAAAAAAATAATCTCCCAGA

250/392

## FIGURE 240

MKITGGLLLLCTVVYFCSSSEAASLSPKKVDCSIYKKYPVVAIPCPITYLPVCGSDYITY  
GNECHLCTESLKSNQGRVQFLHDGSC

Signal peptide:  
amino acids 1-19

251/392

**FIGURE 241**

CCCGCGCCCGGTTCTCCCTCGCAGCACCTCGAAGTGCGCCCCCTCGCCCTCCTGCTCGCGC  
CCCGCCGCCATGGCTGCCTCCCCGCGCGGCCTGCTGTCTGGCCCTGACCGGGCTGGCG  
CTGCTCCTGCTCCTGTGCTGGGGCCCAGGTGGCATAAGTGGAATAAACTCAAGCTGATG  
CTTCAAAAACGAGAAGCACCTGTTCCAATAAGACTAAAGTGCCGTTGATGAGAATAAA  
GCCAAGAATTCTTTGGCAGCCTGAAGCGCCAGAAGCGGCAGCTGTGGGACCGGACTCGG  
CCCGAGGTGCAGCAGTGGTACCAGCAGTTTCTCTACATGGGCTTTGATGAAGCGAAATTT  
GAAGATGACATCACCTATTGGCTTAACAGAGATCGAAATGGACATGAATACTATGGCGAT  
TACTACCAACGTCACCTATGATGAAGACTCTGCAATTGGTCCCCGAGCCCCTACGGCTTT  
AGGCATGGAGCCAGCGTCAACTACGATGACTACTAACCATGACTTGCCACACGCTGTACA  
AGAAGCAAATAGCGATTCTCTTCATGTATCTCCTAATGCCTTACACTACTTGGTTTCTGA  
TTTGCTCTATTTGAGCAGATCTTTTCTACCTACTTTGTGTGATCAAAAAGAAGAGTTAA  
ACAACACATGTAAATGCCTTTTGATATTTTCATGGGAATGCCTCTCATTTAAAAATAGAA  
ATAAAGCATTTTGTAAAAAGA

252/392

## FIGURE 242

MAASPARPAVLALTGLALLLLLCWGPGGISGNKLLMLQKREAPVPTKTKVAVDENKAKE  
FLGSLKRQKRQLWDRTRPEVQQWYQQFLYMGFDEAKFEDDITYWLNDRDRNGHEYYGDDYYQ  
RHYDEDSAIGPRSPYGFRHGASVNYDDY

Signal peptide:  
amino acids 1-30

253/392

## FIGURE 243

CTCCATTAAACCACCACCAGCTCCCCAAGCCACCCCTTCAGCCATGAAGTTCCTGCTCCT  
GGTCTTGGCAGCCCTCGGATTCTTGACCCAGGTGATCCCAGCCAGTGCAGGTGGGTCAAA  
ATGTGTGAGTAACACCCAGGATACTGCAGGACATGTTGCCACTGGGGGGAGACAGCATT  
GTTTCATGTGCAACGCTTCCAGAAAATGCTGCATCAGCTACTCCTTCCTGCCGAAGCCTGA  
CCTACCACAGCTCATCGGTAACCACTGGCAATCAAGGAGAAGAAACACACAAAGGAAAGA  
CAAGAAGCAACAAACGACCGTAACATCATAAATAACCACTGCTATCGCCTCCACCAACTCA  
GAGAAATATCATTTCCACAGTTCCAATTCCTCCTACATTGCTGAGTACTAGCCAAGGCTC  
CTCTTTATGGGGCAGATATCTATAGCCAACCCCAAACTTCTGTCTTCTATCATTCTGTC  
ATTCATCTAGTAATAATTTGGAGTTTGTATCTATCTTACGAGAACAATCATCATGCAGA  
TTCGTCCACAGGGGATCTGTCTAGTTTGGGTCCTCCAAATGAAAAATGTCAAGACAGAATT  
GGACATGCAAAAGATTGACTGGGAGAACACACCTCTGATGGACAAAGGTGAGACAGCA  
GCCACAGGCAGGGAGAGCCTTCAGACTGCAACGCTGGCCTGATACGTGTCAAAGGAGAGA  
GGGATAGAGGAGGATTGAATAGAAGGAGACTAAGACTGCAGCTCTAAGAAAGTCTCAGCC  
AAACAGATGGGGAGGCCCAAAGCAAGGCTTGCCCCCTCAGAGGAGCTCACGCAGGGCAGGA  
ATAGCCAGGTTCTCATATCCCAGGGGTTTCAGACTTGGCTGAGAACAGCCCCCTGGAGAACA  
TGGGGTGACTGCTACCATAGGTCTGGAAGTATGAGGCTGTCCACCAACTATCCCCTTGAA  
GCAAGTTCTCTTGAAAGGAAATCTAAACAGTGCACCCCCATGGCTGCCACGGAGTATAAG  
GAGGGAGAGAAAGGAGCTGAAAGTCTAGGTTTGGCCAGCTAGGTAGACTGACTTGTGAGG  
TATTTATTTATTCATTTGAGTAACAAAGCAGACAGAATACATAGCCACCATTGGTAGTAC  
ACCCCAAAGCAAGGATGGCATGATGCTGGTGACTCAAACGTGCCTACTCATGGTGTCAA  
ATTGGCATAATCCTCTTGGAAGCTGTGTGGAAATAAGCACAGAGAAGCAGAACTCTAAT  
TGCTTAATCCACTAAACATTACTTCTGGGAATTGGCTCATCATAAATTATCCAAGAGAAA  
GCACAAAGTTATGGGCACAAAGGTTTTCCATATAATATTATTTAAATGCTGAGAAAATG  
AAAAATCTAAATGGTGAAATATATACTAATGCCATCTATAAATACAAACAAATAGAATG  
TTTATAGAATAATGGAACATAATAACATTATTCAAAATTGCATTTATGCTATAGTTGTCA  
AAATTGTCTCCTTATATGATACAAAACCTCATGAAAATTATGACTTTTTTGTGTTGGTGG  
AAGCAGAATTATGCATAAATTTCTCTTACAGTTTCGATGCCCATTAGTTTATATAACAT  
TTATTTGACACGTACTGACTTCTATCTGAGAAGAACAACCAAAACACTCAGGCCTAAAT  
AATTA AAAACGGTCC TAAAACTAGCAAACCAGATAAGAAAAGATGTTAATGCCCATTTCC  
CTAACTTATGTCTTAGACCAAATTAATTCTAGATGGTTTTAAATGACAGTGTAAGT  
AAAGTATTAAGATTTGTGTGGTCAAATATTCAATTTAAGAGCAAGGAAATCTTATAAA  
TATAACAATAGAGGCAGAACTCATGTAAGAATAAATTGATTAGGTGGTATTAAATATTAA  
GTTCTTATGTATGTCAAAGATATCATTTTGAAATTCATCCATCTTATTGGGTATTGCAG  
GAGTTCATTCTTTTTGTTTATAAATACTCTCCGTCATATGAATAGTATTCAATTTGTAT  
ACTGGTTTGTGATGGACATTTGGGTTGTTCCAGTTTATGGCTATTACAAATAAGCTT  
CTATGAACATTTATGTACA

254/392

## FIGURE 244

MKFLLLVLAALGFLTQVIPASAGGSKCVSNTPGYCRTCCHWGETALFMCNASRKCCISYS  
FLPKPDLPLIGNHWQSRRTQQRKDKKQQTTVTS

Important features of the protein:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 1-22

N-glycosylation site:

amino acids 50-53

cAMP- and cGMP-dependent protein kinase phosphorylation site:

amino acids 79-82

N-myristoylation site:

amino acids 23-28

255/392

**FIGURE 245**

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGG  
CGGAGCCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTG  
CCCGGCAGCCGGGAGCCATGCGACCCCAGGGCCCCGCGCCTCCCCGCAGCGGCTCCGCG  
GCCTCCTGCTGCTCCTGCTGCTGCAGCTGCCCCGCGCGTTCGAGCGCCTCTGAGATCCCCA  
AGGGGAAGCAAAAGGCGCAGCTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGT  
GCTTACAAGGGCCAGCAGGAGTGCTGGTTCGAGACGGGAGCCCTGGGGCCAATGTTATTC  
CGGGTACACCTGGGATCCCAGGTGCGGATGGATTCAAAGGAGAAAAGGGGGAATGTCTGA  
GGGAAAGCTTTGAGGAGTCCTGGACACCCAACTACAAGCAGTGTTTATGGAGTTCATTGA  
ATTATGGCATAGATCTTGGGAAAATTGCGGAGTGACATTTACAAAGATGCGTTCAAATA  
GTGCTCTAAGAGTTTTGTTTCAAGTGGCTCACTTCGGCTAAAATGCAGAAATGCATGCTGTC  
AGCGTTGGTATTTTACATTCAATGGAGCTGAATGTTTCAAGACCTCTTCCCATTTGAAGCTA  
TAATTTATTTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAATATTCATCGCACTT  
CTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGGCTGGATTAGTGGATGTTGCTATCTGGG  
TTGGCACTTGTTTCAAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTC  
GCATCATTATTGAAGAATAACCAAAATAAATGCTTTAATTTTCAATTTGCTACCTCTTTTT  
TTATTATGCCTTGGAATGGTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCT  
GAATGAAAAGCAAAGCTAAATATGTTTACAGACCAAAGTGTGATTTTCACTGTTTTTAA  
ATCTAGCATTATTCAATTTGCTTCAATCAAAGTGGTTTCAATATTTTTTTTAGTTGGTT  
AGAATACTTTCTTCATAGTCACATTCTCTCAACCTATAATTTGGAATATTGTTGTGGTCT  
TTTGTTTTTTCTCTTAGTATAGCATTTTAAAAAAATATAAAAGCTACCAATCTTTGTAC  
AATTTGTAAATGTTAAGAATTTTTTTTTATATCTGTAAATAAAAATTATTTCCAACA

256/392

## FIGURE 246

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPA  
GVPGRDGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDL  
GKIAECTFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGGLPIEAIYLDQ  
GSPEMNSTINIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIEEE  
LPK

**Signal peptide:**  
amino acids 1-30

**Transmembrane domain:**  
amino acids 195-217



257/392

**FIGURE 247**

GGCCGTTGGTTGGTGC GCGGCTGAAGGGTGTGGCGGAGCAGCGTCTGTTGGTTGGCCGGC  
GGCGGGCCGGGACGGGCATGGCCCTGCTGCTGTGCCTGGTGTGCCTGACGGCGGCGCTGG  
CCCACGGCTGTCTGCACTGCCACAGCAACTTCTCCAAGAAGTTCTCCTTCTACCGCCACC  
ATGTGAACTTCAAGTCCTGGTGGGTGGGCGACATCCCCGTGTCAGGGGCGCTGCTCACCG  
ACTGGAGCGACGACACGATGAAGGAGCTGCACCTGGCCATCCCCGCCAAGATCACCCGGG  
AGAAGCTGGACCAAGTGGCGACAGCAGTGTACCAGATGATGGATCAGCTGTACCAGGGGA  
AGATGTACTTCCCCGGGTATTTCCCCAACGAGCTGCGAAACATCTTCCGGGAGCAGGTGC  
ACCTCATCCAGAACGCCATCATCGAAAGGCACCTGGCACCAGGCAGCTGGGGAGGAGGGC  
AGCTCTCCAGGGAGGGACCCAGCCTAGCACCTGAAGGATCAATGCCATCACCCCGCGGGG  
ACCTCCCCTAAGTAGCCCCCAGAGGCGCTGGGAGTGTGCCACCGCCCTCCCCTGAAGTT  
TGCTCCATCTCACGCTGGGGGTCAACCTGGGGACCCCTTCCCTCCGGGCCATGGACACAC  
ATACATGAAAACAGGCCGCATCGACTGTCAGCACCGCTGTGGCATCTTCCAGTACGAGA  
CCATCTCCTGCAACAACATGCACAGACTCGCACGTGCGCTGCTTTGGCTATAACTGCGAGT  
AGGGCTCAGGCATCACACCCACCCGTCGCGAGGGCCCTACTGTCCCTGGGGTCCCAGGCTC  
TCCTTGAGGGGGCTCCCCGCTTCCACCTGGCTGTCATCGGGTAGGGCGGGGCCGTGGG  
TTCAGGGGCGCACCACTTCCAAGCCTGTGTCCACAGGTCCCTCGGCGCATGGGAAGTCAG  
CTGTCCAGGGCCTCCTGAACTACATAAATAACTGGCACAAAGTAAGTCCCTCCTCAAACC  
AACACAGGCAGTGTGTGTATGTGAGCACCTCGTGGGTGAGTATGTGTGGGGCACAGGCTG  
GCTCCCTCAGCTCCACGTCTTAGAGGGGCTCCCGAGGAGGTGGAACCTCAACCCAGCTC  
TGCGCAGGAGGCGGCTGCAGTCTTTTCTCCCTCAAAGGTCTCCGACCCCTCAGCTGGAGG  
CGGGCATCTTTCCTAAAGGGTCCCCATAGGGTCTGGTTCCACCCCATCCAGGTCTGTGG  
TCAGAGCCTGGGAGGGTTCCCTACGATGGTTAGGGGTGCCCCATGGAGGGGCTGACTGCC  
CCACATTGCCTTTCAGACAGGACACGAGCATGAGGTAAGGCCGCCCTGACCTGGACTTCA  
GGGGGAGGGGGTAAAGGGAGAGAGGAGGGGGGCTAGGGGGTCTCTAGATCAGTGGGGGC  
ACTGCAGGTGGGGCTCTCCCTATACCTGGGACACCTGCTGGATGTACCTCTGCAACCAC  
ACCCATGTGGTGGTTTCATGAACAGACCAGCTCCTCTGCCTTCTCCTGGCCTGGGACAC  
ACAGAGCCACCCCGGCCTTGTGAGTGACCCAGAGAAGGGAGGCCTCGGGAGAAGGGGTGC  
TCGTAAGCCAACACCAGCGTGCCGCGGCCTGCACACCCCTTCGGACATCCCAGGCACGAGG  
GTGTCTGTGGATGTGGCCACACATAGGACCACACGTCCCAGCTGGGAGGAGAGGCCTGGGG  
CCCCCAGGGAGGGAGGCAGGGGGTGGGGGACATGGAGAGCTGAGGCAGCCTCGTCTCCCC  
GCAGCCTGGTATCGCCAGCCTTAAGGTGTCTGGAGCCCCCACACTTGGCCAACCTGACCT  
TGGAAGATGCTGCTGAGTGTCTCAAGCAGCACTGACAGCAGCTGGGCCTGCCCCAGGGCA  
ACGTGGGGGCGGAGACTCAGCTGGACAGCCCCTGCCTGTCACTCTGGAGCTGGGCTGCTG  
CTGCCTCAGGACCCCTCTCCGACCCCGACAGAGCTGAGCTGGCCAGGGCCAGGAGGGC  
GGGAGGGAGGGAATGGGGTGGGCTGTGCGCAGCATCAGCGCCTGGGCAGGTCCGCAGAG  
CTGCGGGATGTGATTAAAGTCCCTGATGTTTCTC

258/392

## FIGURE 248

MALLLCLVCLTAALAHGCLHCHSNFSKKFSFYRHHVNFKSWVVDIPVSGALLTDWSDDT  
MKELHLAIPAKITREKLDQVATAVYQMMDQLYQGKMYFPGYFPNELRNIFREQVHLIQNA  
IIERHLAPGSWGGGQLSREGPSLAPEGSMPSPRGDLP

Signal peptide:  
amino acids 1-15

259/392

## FIGURE 249

CGACGATGCTACGCGCGCCGGCTGCCTCCTCCGGACCTCCGTAGCGCCTGCCCGCGGCCCT  
TGGCTGCGGCGCTGCTCTCGTCGCTTGGCGCGTGTCTCTTCTAGAGCCGAGGGACCCGG  
TGGCCTCGTCGCTCAGCCCCCTATTTCCGCACCAAGACTCGCTACGAGGATGTCAACCCCG  
TGCTATTGTGCGGCCCCGAGGCTCCGTGGCGGGACCTTGAGCTGCTGGAGGGGACCTGCA  
CCCCGGTGCAGCTGGTGCCTCTATTTCGCCACGGCACCCGCTACCCACCGGTCAAACAGA  
TCCGCAAGCTGAGGCAGCTGCACGGGTGCTGCAGGCCCGCGGGTCCAGGGATGGCGGGG  
CTAGTAGTACCGGCAGCCGCGACCTGGGTGCAGCGCTGGCCGACTGGCCTTTGTGGTAC  
CGGACTGGATGGACGGGCAGCTAGTAGAGAAGGGACGGCAGGATATGCGACAGCTGGCGC  
TGCGTCTGGCCTCGCTCTTCCCGGCCCTTTTCAGCCGTGAGAACTACGGCCGCTGCGGC  
TCATCACCAGTTCCAAGCACCGCTGCATGGATAGCAGCGCCGCCCTTCCTGCAGGGGCTGT  
GGCAGCACTACCACCTGGCTTGCCGCCGCCGGACGTGCGAGATATGGAGTTTGGACCTC  
CAACAGTTAATGATAAACTAATGAGATTTTTTGATCACTGTGAGAAGTTTTTAACTGAAG  
TAGAAAAAATGCTACAGCTCTTTATCACGTGGAAGCCTTCAAACTGGACCAGAAATGC  
AGAACATTTTAAAAAAGTTGCAGCTACTTTGCAAGTGCCAGTAAATGATTTAAATGCAG  
ATTTAATTTCAAGTAGCTTTTTACCTGTTTACATAGATGATGCAAAGGTATTAGAATATTTAAATGATC  
TCTCTTGGTGTGATGCTTTTTGACATAGATGATGCAAAGGTATTAGAATATTTAAATGATC  
TGAAACAATATTGAAAAAGGAGATATGGGTATACATTAAACAGTCGATCCAGCTGCACCT  
TGTTTTCAGGATATCTTTCAGCACTTGGACAAAGCAGTTGAACAGACAAACAAGGTCTCAGC  
CAATTTCTTCTCCAGTCATCCTCCAGTTTGGTTCATGAGAGACTCTTCTTCCACTGCTTT  
CTCTCATGGGCTACTTCAAAGACAAGGAACCCCTAACAGCGTACAATTACAAAAACAAA  
TGCATCGGAAGTTCCGAAGTGGTCTCATTGTACCTTATGCCTCGAACCTGATATTTGTGC  
TTTACCCTGTGAAAATGCTAAGACTCCTAAAGAACAAATCCGAGTGCAGATGTTATTAA  
ATGAAAAGGTGTTACCTTTGGCTTACTACAAGAACTGTTTCATTTTATGAAGATCTGA  
AGAACCCTACAAGGACATCCTTCAGAGTTGTCAAACCAGTGAAGAATGTGAATTAGCAA  
GGGCTAACAGTACATCTGATGAACATATGAGTAAGTGAAGAATTTTTAATTCTTTAGGA  
ATCTGCAATGAGTGATTACATGCTTGTAATAGGTAGGCAATTCCTTGATTACAGGAAGCT  
TTTATATTACTTGAGTATTTCTGTCTTTTACAGAAAAACATTGGGTTTCTCTCTGGGT  
TGGACATGAAATGTAAGAAAAGATTTTTCACTGGAGCAGCTCTCTTAAGGAGAAACAAAT  
CTATTTAGAGAAACAGCTGGCCCTGCAAATGTTTACAGAAATGAAATTCCTCCTACTTAT  
ATAAGAAATCTCACACTGAGATAGAATTGTGATTTTATAATAACACTTGAAAAGTGCTGG  
AGTAACAAAATATCTCAGTTGGACCATCCTTAACTTGATTGAACTGTCTAGGAACCTTAC  
AGATTGTTCTGCAGTTCTCTCTTCTTTTCTCAGGTAGGACAGCTCTAGCATTTTCTTAA  
TCAGGAATATTGTGGTAAGCTGGGAGTATCACTCTGGAAGAAAGTAACATCTCCAGATGA  
GAATTTGAAACAAGAAACAGAGTGTTGTAAAAGGACACCTTCACTGAAGCAAGTCGGAAA  
GTACAATGAAAATAAATATTTTGGTATTTATTTATGAAATATTTGAACATTTTTTCAAT  
AATTCCTTTTTTACTTCTAGGAAGTCTCAAAGACCATCTTAAATTATTATGTTTGGAC  
AATTAGCAACAAGTCAGATAGTTAGAAATCGAAGTTTTCAAATCCATGCTTAGCTAACT  
TTTTTCAATCTGTCACTTGGCTTCGATTTTTTATATTTTCTTATTATGAAATGTATCTTT  
TGGTTGTTTGATTTTTCTTTCTTTCTTTGTAAATAGTTTCTGAGTTCTGTCAAATGCCGTG  
AAAGTATTTGCTATAATAAAGAAAATTCCTGTGACTTTAAAAA

260/392

**FIGURE 250**

MLRAPGCLLRTSVAPAAALAAALLSSLARCSLLEPRDPVASSLSFYFGTKTRYEDVNPVL  
LSGPEAPWRDPELLEGTCTPVQLVALIRHGTRYPTVKQIRKLRQLHGLLQARGSRDGGAS  
STGSRDLGAALADWPLWYADWMDGQLVEKGRQDMRQLALRLASLFPALFSRENYGRLRLI  
TSSKHRCMDSSAAFLOGLWQHYHPGLPPPDVADMEFGPPTVNDKLMRFFDHCEKFLTEVE  
KNATALYHVEAFKTGPEMQNILKKVAATLQVPVNDLNADLIQVAFFTCSFDLAIKGVKSP  
WCDVFDIDDAKVLEYLNDLKQYWKRGYGYTINSRSSCTLFQDIFQHLDKAVEQKQRSQPI  
SSPVILQFGHAETLLPLLMLGYFKDKEPLTAYNYKKQMRKFRSGLIVPYASNLI FVLY  
HCENAKTPKEQFRVQMLLNEKVLPLAYSQETVSFYEDLKNHYKDILQSCQTSEECELARA  
NSTSDEL

**Important features:****Signal sequence**

amino acids 1-30

**N-glycosylation sites:**

amino acids 242-246, 481-485

**N-myristoylation sites.**

amino acids 107-113, 113-119, 117-123, 118-124, 128-134

**Endoplasmic reticulum targeting sequence:**

amino acids 484-489

261/392

**FIGURE 251**

GGAGAGCCGCGGCTGGGACCGGAGTGGGGAGCGCGGCGTGGAGGTGCCACCCGGCGCGGG  
TGGCGGAGAGATCAGAAGCCTCTTCCCCAAGCCGAGCCAACCTCAGCGGGGACCCGGGCT  
CAGGGACGCGGCGGCGGCGGCGGCGACTGCAGTGGCTGGACGATGGCCAGCGTCCGCCGGA  
GCCGGGGCGGTGATTGCAGCCCCAGACAGCCGCGCGCTGGCTGTGGTTCGGTGTCTGGCGGCG  
GCGCTTGGGCTCTTGACAGCTGGAGTATCAGCCTTGGAAAGTATATACGCCAAAAGAAATC  
TTCGTGGCAAATGGTACACAAGGGAAGCTGACCTGCAAGTTCAAGTCTACTAGTACGACT  
GGCGGGTTGACCTCAGTCTCCTGGAGCTTCCAGCCAGAGGGGGCCGACACTACTGTGTCTG  
TTTTTCCACTACTCCCAAGGGCAAGTGTACCTTGGGAATTATCCACCATTAAAGACAGA  
ATCAGCTGGGCTGGAGACCTTGACAAGAAAGATGCATCAATCAACATAGAAAATATGCAG  
TTTATACACAATGGCACCTATATCTGTGATGTCAAAAACCCTCCTGACATCGTTGTCCAG  
CCTGGACACATTAGGCTCTATGTCTGTAGAAAAAGAGAATTTGCCTGTGTTTCCAGTTTGG  
GTAGTGGTGGGCATAGTTACTGCTGTGGTCTTAGGTCCTCACTCTGCTCATCAGCATGATT  
CTGGCTGTCTCTATAGAAGGAAAACTCTAAAACGGGATTACACTGGCTGCAGTACATCA  
GAGAGTTTGTCAACAGTTAAGCAGGCTCCTCGGAAGTCCCCCTCCGACACTGAGGGTCTT  
GTAAAGAGTCTGCCTTCTGGATCTCACCAGGGCCCAGTCATATATGCACAGTTAGACCAC  
TCCGGCGGACATCACAGTGACAAGATTAAACAAGTCAGAGTCTGTGGTGTATGCGGATATC  
CGAAAGAATTAAGAGAATACCTAGAACATATCCTCAGCAAGAAACAAAACCAAACCTGGAC  
TCTCGTGCAGAAAATGTAGCCCATTAACCATGTAGCCTTGGAGACCCAGGCAAGGACAA  
GTACACGTGTACTCACAGAGGGAGAGAAAGATGTGTACAAAGGATATGTATAAATATTCT  
ATTTAGTCATCCTGATATGAGGAGCCAGTGTTCATGATGAAAAGATGGTATGATTCTAC  
ATATGTACCCATTGTCTTGCTGTTTTTGTACTTTCTTTTCAGGTCATTTACAATTGGGAG  
ATTTCAGAAACATTCTTTTACCATCATTTAGAAATGGTTTGCCTTAATGGAGACAATAG  
CAGATCCTGTAGTATTTCCAGTAGACATGGCCTTTTAATCTAAGGGCTTAAGACTGATTA  
GTCTTAGCATTTTACTGTAGTGGAGGATGGAGATGCTATGATGGAAGCATACCCAGGGTG  
GCCTTTAGCCAGATATCAGTACCATTTTATTGTCTGCCGCTTTTAAAAAATACCCATTGG  
CTATGCCACTTGAAAAACAATTTGAGAAGTTTTTTTTTGAAGTTTTTCTCACTAAAATATGGG  
GCAATTGTTAGCCTTACATGTTGTGTAGACTTACTTTAAGTTTGACCCCTTGAATGTGT  
CATATCAATTTCTGGATTATAATAGCAAGATTAGCAAAGGATAAATGCCGAAGGTCACT  
TCATTCTGGACACAGTTGGATCAATACTGATTAAAGTAGAAAATCCAAGCTTTGCTTGAGA  
ACTTTTGTAAACGTGGAGAGTAAAAAGTATCGGTTTTTA

262/392

## FIGURE 252

MAASAGAGAVIAAPDSRRWLWSVLAAALGLLTAGVSALEVYTPKEIFVANGTQGKLTCKF  
KSTSTTGGLTSVSWSFQPEGADTTVSFFHYSQGQVYLGNYPPFKDRISWAGDLDDKIDASI  
NIENMQFIHNGTYICDVKNPPDIVVQPGHIRLYVVEKENLPVFPVWVVVGIVTAVVLGLT  
LLISMILAVLYRRKNSKRDTGTCSTSESLSPVKQAPRKSPSDTEGLVKSLPSGSHQGPVI  
YAQLDHSGGHHSDKINKSESVVYADIRKN

Signal peptide:  
amino acids 1-37

Transmembrane domain:  
amino acids 161-183

263/392

**FIGURE 253**

GTGACACTATAGAAGAGCTATGACGTCGCATGCACGCGTACGTAAGCTCGGAATTCGGCT  
CGAGGCTGGTGGGAAGAAGCCGAGATGGCGGCAGCCAGCGCTGGGGCAACCCGGCTGCTC  
CTGCTCTTGCTGATGGCGGTAGCAGCGCCAGTCGAGCCCGGGGCAGCGGCTGCCGGGCC  
GGGACTGGTGC GCGAGGGGCTGGGGCGGAAGGTCGAGAGGGCGAGGCCTGTGGCACGGTG  
GGGCTGCTGCTGGAGCACTCATTTGAGATCGATGACAGTGCCAACTTCCGGAAGCGGGGC  
TCACTGCTCTGGAACCAGCAGGATGGTACCTTGTCCCTGTCACAGCGGCAGCTCAGCGAG  
GAGGAGCGGGCCGACTCCGGGATGTGGCAGCCCTGAATGGCCTGTACCGGGTCCGGATC  
CCAAGGCGACCCGGGGCCCTGGATGGCCTGGAAGCTGGTGGCTATGTCTCCTCCTTTGTC  
CCTGCGTGCTCCCTGGTGGAGTCGCACCTGTGCGACCAGCTGACCCTGCACGTGGATGTG  
GCCGGCAACGTGGTGGGCGTGTCGGTGGTGACGCACCCCGGGGGCTGCCGGGGCCATGAG  
GTGGAGGACGTGGACCTGGAGCTGTTCAACACCTCGGTGCAGCTGCAGCCGCCACCACA  
GCCCCAGGCCCTGAGACGGCGGCCTTCATTGAGCGCCTGGAGATGGAACAGGCCCCAGAAG  
GCCAAGAACCCCCAGGAGCAGAAGTCCTTCTTCGCCAAATACTGGATGTACATCATTCCTC  
GTCGTCCTGTTCTCATGATGTGAGGAGCGCCAGACACCGGGGGCCAGGGTGGGGGTGGG  
GGTGGGGGTGGTGGTGGGGGTAGTGGCCTTTGCTGTGTGCCACCCTCCCTGTAAGTCTAT  
TTAAAAACATCGACGATACATTGAAATGTGTGAACGTTTTGAAAAGCTACAGCTTCCAGC  
AGCCAAAAGCAACTGTTGTTTTGGCAAGACGGTCCTGATGTACAAGCTTGATTGAAATTC  
ACTGCTCACTTGATACGTTATTCAGAAACCCAAGGAATGGCTGTCCCCATCCTCATGTGG  
CTGTGTGGAGCTCAGCTGTGTTGTGTGGCAGTTTATTAACTGTCCCCAGATCGACACG  
CAAAAAAAAA

264/392

**FIGURE 254**

MAAASAGATRLLLLLLLMAVAAPSRARGSGCRAGTGARGAGAEGREGEACGTVGLLLEHSF  
EIDDSANFRKRGSLLWNQQDGTLSLSQRQLSEEERGRIRDVAALNGLYRVRI PRRPGALD  
GLEAGGYVSSFVPACSLVESHLSQDLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDLEL  
FNTSVQLQPPTTAPGPETA AFIERLEMEQAQKAKNPQEQKSFFAKYWYIIPVVLFLMMS  
GAPDTGGQGGGGGGGGGGGSGLCCVPPSL

Signal peptide:  
amino acids 1-24

Transmembrane domain:  
amino acids 226-243



265/392

**FIGURE 255**

GCGACGCGCGGCGGGGCGGCGAGAGGAAACGCGGCGCCGGGCCGGGCCCGGCCCTGGAGA  
TGGTCCCCGCGCGCCGCGGGCTGGTGTTGTCTCGTGCTCTGGCTCCCCGCGTGCGTCGCGG  
CCCACGGCTTCCGTATCCATGATTATTTGTACTTTCAAGTGCTGAGTCCTGGGGACATTC  
GATACATCTTCACAGCCACACCTGCCAAGGACTTTGGTGGTATCTTTCACACAAGGTATG  
AGCAGATTCACCTTGTCCCCGCTGAACCTCCAGAGGCCTGCGGGGAACCTCAGCAACGGTT  
TCTTCATCCAGGACCAGATTGCTCTGGTGGAGAGGGGGGGCTGCTCCTTCCTCTCCAAGA  
CTCGGGTGGTCCAGGAGCACGGCGGGCGGGCGGTGATCATCTCTGACAACGCAGTTGACA  
ATGACAGCTTCTACGTGGAGATGATCCAGGACAGTACCCAGCGCACAGCTGACATCCCCG  
CCCTCTTCCTGCTCGGCCGAGACGGCTACATGATCCGCCGCTCTCTGGAACAGCATGGGC  
TGCCATGGGCCATCATTCCATCCCAGTCAATGTCACCAGCATCCCCACCTTTGAGCTGC  
TGCAACCGCCCTGGACCTTCTGGTAGAAGAGTTTGTCCCACATTCAGCCATAAGTGA  
CTGAGCTGGGAAGGGGAAACCCAGGAATTTTGCTACTTGGAATTTGGAGATAGCATCTGG  
GGACAAGTGGAGCCAGGTAGAGGAAAAGGTTTGGGCGTTGCTAGGCTGAAAGGGAAGCC  
ACACCACTGGCCTTCCCTTCCCCAGGGCCCCCAAGGTGTCTCATGCTACAAGAAGAGGC  
AAGAGACAGGCCCCAGGGCTTCTGGCTAGAACCCGAAACAAAAGGAGCTGAAGGCAGGTG  
GCCTGAGAGCCATCTGTGACCTGTCACACTCACCTGGCTCCAGCCTCCCCTACCCAGGGT  
CTCTGCACAGTGACCTTCACAGCAGTTGTTGGAGTGGTTTAAAGAGCTGGTGTTTGGGGA  
CTCAATAAACCTCACTGACTTTTTAGCAATAAAGCTTCTCATCAGGGTTGCAAAAAAA  
AAAAAAAAAAAAAAAAAAAA

266/392

## FIGURE 256

MVPGAAGWCCLVLWLPACVAAHGFRIDYLYFQVLSPGDIRYIFTATPAKDFGGIFHTRY  
EQIHLVPAEPPEACGELSNGFFIQDQIALVERGGCSFLSKTRVVQEHGGRAVIISDNAV  
NDSFYVEMIQDSTQRTADIPALFLLGRDGYMIRRSLEQHGLPWAIISIPVNVTSIPTFEL  
LQPPWTFW

Signal peptide:  
amino acids 1-20

267/392

**FIGURE 257**

CTCGCTTCTTCCTTCTGGATGGGGGCCCAGGGGGCCCAGGAGAGTATAAAGGCGATGTGG  
AGGGTGCCCGGCACAACCAGACGCCCAGTCACAGGCGAGAGCCCTGGGATGCACCGGCCA  
GAGGCCATGCTGCTGCTGCTCACGCTTGCCCTCCTGGGGGGCCCCACCTGGGCAGGGAAG  
ATGTATGGCCCTGGAGGAGGCAAGTATTTTCAGCACCACTGAAGACTACGACCATGAAATC  
ACAGGGCTGCGGGTGTCTGTAGGTCTTCTCCTGGTGAAAAGTGTCCAGGTGAAACTTGGA  
GACTCCTGGGACGTGAAACTGGGAGCCTTAGGTGGGAATACCCAGGAAGTCACCCCTGCAG  
CCAGGCGAATACATCACAAAAGTCTTTGTGCGCCTTCCAAGCTTTCCTCCGGGGTATGGTC  
ATGTACACCAGCAAGGACCGCTATTTCTATTTTGGGAAGCTTGATGGCCAGATCTCCTCT  
GCCTACCCCAGCCAAGAGGGGCAGGTGCTGGTGGGCATCTATGGCCAGTATCAACTCCTT  
GGCATCAAGAGCATTGGCTTTGAATGGAATTATCCACTAGAGGAGCCGACCACTGAGCCA  
CCAGTTAATCTCACATACTCAGCAAACCTACCCGTGGGTGCTAGGGTGGGGTATGGGGC  
CATCCGAGCTGAGGCCATCTGTGTGGTGGTGGCTGATGGTACTGGAGTAACTGAGTCGGG  
ACGCTGAATCTGAATCCACCAATAAATAAAGCTTCTGCAGAAAA

268/392

## FIGURE 258

MHRPEAML<sup>1</sup>LLTLALLGGPTWAGKMYGPGGGKYFSTTEDYDHEITGLRVSVGLLLVKS<sup>22</sup>VQ  
VKLGDSWDVKLGALGGNTQEVTLPGEYITKVFAFQAFLRGMVMTSKDRYFYFGKLDG  
QISSAYPSQEGQVLVGIYGQYQLLGIKSIGFEWNYPLEEPTTEPPVNLTYSANSPVGR

Signal peptide:  
amino acids 1-22

269/392

**FIGURE 259**

CAGACATGGCTCAGTCACTGGCTCTGAGCCTCCTTATCCTGGTTCTGGCCTTTGGCATCC  
CCAGGACCCAAGGCAGTGATGGAGGGGCTCAGGACTGTTGCCTCAAGTACAGCCAAAGGA  
AGATTCCCGCCAAGGTTGTCCGCAGCTACCGGAAGCAGGAACCAAGCTTAGGCTGCTCCA  
TCCCAGCTATCCTGTTCTTGCCCCGCAAGCGCTCTCAGGCAGAGCTATGTGCAGACCCAA  
AGGAGCTCTGGGTGCAGCAGCTGATGCAGCATCTGGACAAGACACCATCCCCACAGAAAC  
CAGCCCAGGGCTGCAGGAAGGACAGGGGGGCCTCCAAGACTGGCAAGAAAGGAAAGGGCT  
CCAAAGGCTGCAAGAGGACTGAGCGGTCACAGACCCCTAAAGGGCCATAGCCCAGTGAGC  
AGCCTGGAGCCCTGGAGACCCACACAGCCTCACCAGCGCTTGAAGCCTGAACCCAAGATG  
CAAGAAGGAGGCTATGCTCAGGGGCCCTGGAGCAGCCACCCCATGCTGGCCTTGCCACAC  
TCTTTCTCCTGCTTTAACCACCCCATCTGCATTCCCAGCTCTACCCTGCATGGCTGAGCT  
GCCACAGCAGGCCAGGTCCAGAGAGACCGAGGAGGGAGAGTCTCCAGGGAGCATGAGA  
GGAGGCAGCAGGACTGTCCCCTTGAAGGAGAATCATCAGGACCCTGGACCTGATACGGCT  
CCCCAGTACACCCACCTCTTCCTTGTAATATGATTTATACCTAACTGAATAAAAAGCT  
GTTCTGTCTTCCCNCCCA

270/392

## FIGURE 260

MAQSLALSLLILVLAFGIPRTQGS DGG A QDCCLKYSQRKIPAKVVRSYRKQEPSLGCSIP  
AILFLPRKRSQAELCADPKELWVQQLMQHLDKTPSPQKPAQGCRKDRGASKTGKKKGKGSK  
GCKRTERSQTPKGP

**Important features of the protein:**

**Signal peptide:**

amino acids 1-17

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

amino acids 67-71

**N-myristoylation sites:**

amino acids 17-23, 23-29, 27-33, 108-114, 118-124, 121-127

**Amidation site:**

amino acids 112-116

**Small cytokines:**

amino acids 51-91

271/392

## FIGURE 261

GGGACTACAAGCCGCGCCGCGCTGCCGCTGGCCCCCTCAGCAACCCCTCGACATGGCGCTGA  
GGCGGCCACCGCGACTCCGGCTCTGCGCTCGGCTGCCTGACTTCTTCCTGCTGCTGCTTT  
TCAGGGGCTGCCTGATAGGGGCTGTAAATCTCAAATCCAGCAATCGAACCCCACTGGTAC  
AGGAATTTGAAAGTGTGGAACGTCTTGCATCATTACGGATTTCGACAGACAAGTGACCCCA  
GGATCGAGTGGAAGAAAATTCAAGATGAACAAACCACATATGTGTTTTTTTGACAACAAAA  
TTCAGGGGAGACTTGGCGGGTCGTGCAGAAATACTGGGGAAGACATCCCTGAAGATCTGGA  
ATGTGACACGGAGAGACTCAGCCCTTTATCGCTGTGAGGTCGTTGCTCGAAATGACCGCA  
AGGAAATTGATGAGATTGTGATCGAGTTAACTGTGCAAGTGAAGCCAGTGACCCCTGTCT  
GTAGAGTGCCGAAGGCTGTACCAGTAGGCAAGATGGCAACACTGCACTGCCAGGAGAGTG  
AGGGCCACCCCGGCCTCACTACAGCTGGTATCGCAATGATGTACCCTGCCACGGATT  
CCAGAGCCAATCCCAGATTTTCGAATTTCTTCTTTCCACTTAAACTCTGAAACAGGCACCT  
TGGTGTTCACTGCTGTTTCAAGGACGACTCTGGGCAGTACTACTGCATTGCTTCCAATG  
ACGCAGGCTCAGCCAGGTGTGAGGAGCAGGAGATGGAAGTCTATGACCTGAACATTGGCG  
GAATTATTGGGGGGGTTCTGGTTGTCTTGTCTGTACTGGCCCTGATCACGTTGGGCATCT  
GCTGTGCATACAGACGTGGCTACTTCATCAACAATAAACAGGATGGAGAAAGTTACAAGA  
ACCCAGGGAACCCAGATGGAGTTAACTACATCCGCACTGACGAGGAGGGCGACTTCAGAC  
ACAAGTCATCGTTTGTGATCTGAGACCCGCGGTGTGGCTGAGAGCGCACAGAGCGCACGT  
GCACATACCTCTGCTAGAACTCCTGTCAAGGCAGCGAGAGCTGATGCACTCGGACAGAG  
CTAGACACTCATTGAGAAGCTTTTCGTTTTGGCCAAAGTTGACCACTACTCTTCTTACTC  
TAACAAGCCACATGAATAGAAGAAATTTTCTCAAGATGGACCCGGTAAATATAACCACAA  
GGAAGCGAAACTGGGTGCGTTCACTGAGTTGGGTTCTTAATCTGTTTCTGGCCTGATTCC  
CGCATGAGTATTAGGGTGATCTTAAAGAGTTTGCTCACGTAAACGCCCGTGTGGGCCCT  
GTGAAGCCAGCATGTTTACCCTGGTTCGTTGAGCAGCCACGACAGCACCATTGTGAGATGG  
CGAGGTGGCTGGACAGCACAGCAGCGCATCCCGCGGGAACCCAGAAAGGCTTCTTAC  
ACAGCAGCCTTACTTCATCGGCCCCACAGACACCACCGCAGTTTCTTCTTAAAGGCTCTGC  
TGATCGGTGTTGCAGTGTCCATTGTGGAGAAGCTTTTGGATCAGCATTTTGTAAAAACA  
ACCAAAATCAGGAAGGTAAATTGGTTGCTGGAAGAGGGATCTTGCTGAGGAACCCCTGCT  
TGTCCAACAGGGTGTCAGGATTTAAGGAAAACCTTCGTTCTAGGCTAAGTCTGAAATGGT  
ACTGAAATATGCTTTTCTATGGGTCTGTTTATTTTATAAAATTTTACATCTAAATTTTT  
GCTAAGGATGTATTTTGTATTATTGAAAAGAAAATTTCTATTTAACTGTAAATATATTGT  
CATACAATGTTAAATAACCTATTTTTTTTAAAAAAGTTCAACTTAAGGTAGAAGTTCCAAG  
CTACTAGTGTTAAATTGGAAAATATCAATAATTAAGAGTATTTTACCAAGGAATCCTCT  
CATGGAAGTTTACTGTGATGTTCTTTTCTCACACAAGTTTTAGCCTTTTTTACAAGGGA  
ACTCATACTGTCTACACATCAGACCATAGTTGCTTAGGAAACCTTTAAAAATTCAGTTA  
AGCAATGTTGAAATCAGTTTGCATCTCTTCAAAAGAAACCTCTCAGGTAGCTTTGAACT  
GCCTCTTCTGAGATGACTAGGACAGTCTGTACCCAGAGGCCACCCAGAAGCCCTCAGAT  
GTACATACAGATGCCAGTCAGCTCCTGGGGTTGCGCCAGGCGCCCCCGCTCTAGCTCA  
CTGTTGCCTCGCTGTCTGCCAGGAGGCCCTGCCATCCTTGGGCCCTGGCAGTGGCTGTGT  
CCCAGTGAGCTTTACTCAGTGGCCCTTGCTTCATCCAGCACAGCTCTCAGGTGGGCACT  
GCAGGGACACTGGTGTCTTCCATGTAGCGTCCCAGCTTTGGGCTCCTGTAACAGACCTCT  
TTTTGGTTATGGATGGCTCACAAAATAGGGCCCCCAATGCTATTTTTTTTTTTTAAAGTTT  
GTTTAATTATTTGTTAAGATTGTCTAAGGCCAAAGGCAATTGCGAAATCAAGTCTGTCAA  
GTACAATAACATTTTAAAGAAAATGGATCCCCTGTTCTCTTTGCCACAGAGAAAGC  
ACCCAGACGCCACAGGCTCTGTGCGATTTCAAAACAAACCATGATGGAGTGGCGGGCAGT  
CCAGCCTTTTAAAGAACGTCAGGTGGAGCAGGCAGGTGAAAGGCCTGGCGGGGAGGAAAG  
TGAAACGCCTGAATCAAAAGCAGTTTTTCTAATTTTGACTTTAAATTTTTTCATCCGCCGGA

272/392

GACACTGCTCCCATTTGTGGGGGGACATTAGCAACATCACTCAGAAGCCTGTGTTCTTCA  
AGAGCAGGTGTTCTCAGCCTCACATGCCCTGCCGTGCTGGACTCAGGACTGAAGTGCTGT  
AAAGCAAGGAGCTGCTGAGAAGGAGCACTCCACTGTGTGCCTGGAGAATGGCTCTCACTA  
CTCACCTTGTCTTTCAGCTTCCAGTGTCTTGGGTTTTTTTATACTTTGACAGCTTTTTTTTT  
AATTGCATACATGAGACTGTGTTGACTTTTTTTTAGTTATGTGAAACACTTTGCCGCAGGC  
CGCCTGGCAGAGGCAGGAAATGCTCCAGCAGTGGCTCAGTGCTCCCTGGTGTCTGCTGCA  
TGGCATCCTGGATGCTTAGCATGCAAGTTCCCTCCATCATTGCCACCTTGGTAGAGAGGG  
ATGGCTCCCCACCTCAGCGTTGGGGATTACGCTCCAGCCTCCTTCTTGGTTGTCATAG  
TGATAGGGTAGCCTTATTGCCCCCTCTTCTTATACCCTAAAACCTTCTACACTAGTGCCA  
TGGGAACCAGGTCTGAAAAAGTAGAGAGAAGTGAAAGTAGAGTCTGGGAAGTAGCTGCCT  
ATAACTGAGACTAGACGGAAAAGGAATACTCGTGTATTTTAAGATATGAATGTGACTCAA  
GACTCGAGGCCGATACGAGGCTGTGATTCTGCCTTTGGATGGATGTTGCTGTACACAGAT  
GCTACAGACTTGTAATAACACACCGTAATTTGGCATTGTGTTAACCTCATTTATAAAAGC  
TTCAAAAAAACCCA



273/392

**FIGURE 262**

MALRRPPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNRTPVVQEFESVELSCIITDSQT  
SDPRIEWKKIQDEQTTYVFFDNKIQGDLAGRAEILGKTSLSKIWNVTRRDSALYRCEVVAR  
NDRKEIDEIVIELTVQVKPVPVPCRVPKAVPVGKMATLHCQESEGHPRPHYSWYRNDVPL  
PTDSRANPRFRNSSFHLNSETGTLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEVDL  
NIGGIIGGVLVVLAVLALITLGICCAYYRRGYFINNKQDGESYKNPGKPDGVNYIRTDEEG  
DFRHKSSFVI

**Important features of the protein:**

**Signal peptide:**

amino acids 1-30

**Transmembrane domain:**

amino acids 243-263

**N-glycosylation sites:**

amino acids 104-107, 192-195

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

amino acids 107-110

**Casein kinase II phosphorylation site:**

amino acids 106-109, 296-299

**Tyrosine kinase phosphorylation site:**

amino acids 69-77

**N-myristoylation sites:**

amino acids 26-31, 215-220, 226-231, 243-248, 244-249, 262-267

274/392

## FIGURE 263

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGT CAGAGGCCGGGGAAGAGAAGCAA  
AGCGCAACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACC  
CCCTAACTTTCAGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAAGTCCAGCCCCGCACATC  
CACGCGCGGCACAGGCGCGGCAGGCGGCAGGTCCCCGGCCGAAGGCGATGCGCGCAGGGGG  
TCGGGCAGCTGGGCTCGGGCGGCGGGAGTAGGGCCCCGGCAGGGAGGCAGGGAGGCTGCAT  
ATTCAGAGTCGCGGGCTGCGCCCTGGGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCT  
GCTGCCACCGCGCCGCGATGAGCCGCGTGGTCTCGCTGCTGCTGGGCGCCGCGCTGCTCT  
GCGGCCACGGAGCCTTCTGCCGCCGCGTGGTCAGCGGCCAAAAGGTGTGTTTGTGCTGACT  
TCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAAGTGTCCAGCCGAGTGAGCTTTC  
AGGAGGCACGCCTGGCTTGTGAGAGTGAGGGAGGAGTCTCCTCAGCCTTGAGAATGAAG  
CAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGGACAGGGATTT  
CTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAAAACATCTGGTGCCT  
GCCAGATCTCTACCACTGGTCTGATGGAAGCAATTCCAGTACCGAAACTGGTACACAG  
ATGAACCTTCTGCGGAAGTGAAAAGTGTGTTGTGATGTATCACCAACCAACTGCCAATC  
CTGGCCTTGGGGGTCCCTACCTTTACCAGTGGAATGATGACAGGTGTAACATGAAGCACA  
ATTATATTTGCAAGTATGAACCAGAGATTAATCCAACAGCCCCTGTAGAAAAGCCTTATC  
TTACAAATCAACCAGGAGACACCCATCAGAATGTGGTTGTTACTGAAGCAGGTATAATTC  
CCAATCTAATTTATGTTGTTATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTT  
TTGGAACCTGTTGTTTCCAGATGCTGCATAAAAGTAAAGGAAGAACAAAACTAGTCCAA  
ACCAGTCTACACTGTGGATTTCAAAGAGTACCAGAAAAGAAAGTGGCATGGAAGTATAAAT  
AACTCATTTGACTTGGTTCCAGAATTTTGTAAATCTGGATCTGTATAAGGAATGGCATCAG  
AACAATAGCTTGAATGGCTTGAAATCACAAAGGATCTGCAAGATGAAGTGAAGCTCCC  
CCTTGAGGCAAATATTAAAGTAATTTTATATGTCTATTATTTTCAATTTAAAGAATATGCT  
GTGCTAATAATGGAGTGAGACATGCTTATTTTGTCTAAAGGATGCACCCAACTTCAAAT  
TCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTGCGGAGTATGTGT  
GTTAGAAGCAATTCCTTTTATTTCTTTACCTTTCATAAGTTGTTATCTAGTCAATGTAA  
TGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGTTGA  
TAAAAATGAAGTGTCTAATATTTTATTTTATGGCATCTCATTTTTCAATACATGCTCTT  
TTGATTAAAGAACTTATTACTGTTGTCAACTGAATTCACACACACACAAATATAGTACC  
ATAGAAAAAGTTTGTCTTCGAAATAATTCATCTTTCAGCTTCTCTGCTTTTGGTCAAT  
GTCTAGGAAATCTCTTCAGAAATAAGAAGCTATTTCAATTAAGTGTGATATAAACCTCCTC  
AAACATTTTACTTAGAGGCAAGGATTGTCTAATTTCAATTGTGCAAGACATGTGCCTTAT  
AATTATTTTACTTAAATTTAAACAGATTTTGTAAATAATGTAACCTTTGTTAATAGGTGC  
ATAAACACTAATGCAGTCAATTTGAACAAAAGAAGTGACATACACAATATAAATCATATG  
TCTTCACACGTTGCCTATATAATGAGAAGCAGCTCTCTGAGGGTTCTGAAATCAATGTGG  
TCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTTCGGGGTTTGGGATTGACACTGGAGGC  
AGATAGTTGCAAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTGACTATATTAG  
TATACAAAGAGGT CATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAGACAAG  
CACAGCACACAGACATTTTAGGAAGGAAAGGAAGTACGAAATCGTGTGAAAATGGGTGG  
AACCCATCAGTGATCGCATATTCATTGATGAGGGTTTGTCTGAGATAGAAAATGGTGGCT  
CCTTTCTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCTTGTCTTCTCAAGAGAAAG  
TTGTAAGTCTCTGGTCTTCATATGTCCCTGTGCTCCTTTTAACCAATAAAGAGTTCCTG  
TTTCTGGGGGAAA

275/392

**FIGURE 264**

MSRVVSLLLGAALLCGHGAFRRVVSQGKVCFAFKHPCYKMAFHELSRVSFQEARLA  
CESEGGVLLSLENAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDQTSGACPDLYQ  
WSDGSNSQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKY  
EPEINPTAPVEKPYLTNQPGDTHQNVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCF  
QMLHKSKGRTKTSPNQSTLWISKSTRKESGMEV

**Important features of the protein:**

**Signal peptide:**

amino acids 1-21

**Transmembrane domain:**

amino acids 214-235

**N-glycosylation sites:**

amino acids 86-89 and 255-258

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

amino acids 266-269

**N-myristoylation sites:**

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145 and 212-217

276/392

## FIGURE 265

GGAGAATGGAGAGAGCAGTGAGAGTGGAGTCCGGGGTCCTGGTCGGGGTGGTCTGTCTGC  
TCCTGGCATGCCCTGCCACAGCCACTGGGCCCGAAGTTGCTCAGCCTGAAGTAGACACCA  
CCCTGGGTCGTGTGCGAGGCCGGCAGGTGGGCGTGAAGGGCACAGACCGCCTTGTGAATG  
TCTTTCTGGGCATTCCATTTGCCAGCCGCCACTGGGCCCTGACCGGTTCTCAGCCCCAC  
ACCCAGCACAGCCCTGGGAGGGTGTGCGGGATGCCAGCACTGCGCCCCAATGTGCCTAC  
AAGACGTGGAGAGCATGAACAGCAGCAGATTTGTCTCAACGGAAAACAGCAGATCTTCT  
CCGTTTCAGAGGACTGCCTGGTCTCAACGTCTATAGCCCAGCTGAGGTCCCCGAGGGT  
CCGGTAGGCCGGTCATGGTATGGGTCCATGGAGGCGCTCTGATAACTGGCGCTGCCACCT  
CCTACGATGGATCAGCTCTGGCTGCCATGAGGGATGTGGTCTGGTTACAGTCCAGTACC  
GCCTTGGGGTCCTTGGCTTCTTCAAGCACTGGAGATGAGCATGCACCTGGCAACCAGGGCT  
TCCTAGATGTGGTAGCTGCTTTGCGCTGGGTGCAAGAAAACATCGCCCCCTTCGGGGGTG  
ACCTCAACTGTGTCACTGTCTTTGGTGGATCTGCCGGTGGGAGCATCATCTCTGGCCTGG  
TCCTGTCCCCAGTGGCTGCAGGGCTGTTCCACAGAGCCATCACACAGAGTGGGGTCATCA  
CCACCCCAGGGATCATCGACTCTCACCCCTTGGCCCCTAGCTCAGAAAATCGCAAACACCT  
TGGCCTGCAGCTCCAGCTCCCCGGCTGAGATGGTGCAGTGCCTTCAGCAGAAAAGAAGGAG  
AAGAGCTGGTCCTTAGCAAGAAGCTGAAAAATACTATCTATCCTCTCACCGTTGATGGCA  
CTGTCTTCCCCAAAAGCCCCAAGGAATCCTGAAGGAGAAGCCCTTCCACTCTGTGCCCT  
TCCTCATGGGTGTCAACAACCATGAGTTCAGCTGGCTCATCCCCAGGGGCTGGGGTCTCC  
TGGATACAATGGAGCAGATGAGCCGGGAGGACATGCTGGCCATCTCAACACCCGTCTTGA  
CCAGTCTGGATGTGCCCCCTGAGATGATGCCACCCGTATAGATGAATACCTAGGAAGCA  
ATCGGACGCACAAGCCAAATGCCAGGCGTTCAGGAATTCATGGGTGACGTATTTCATCA  
ATGTTCCACCGTCAGTTTTTTCAAGATACCTTCGAGATTCTGGAAGCCCTGTCTTTTTCT  
ATGAGTTCCAGCATCGACCCAGTTCTTTTGCAGAGATCAAACCTGCCTGGGTGAAGGCTG  
ATCATGGGGCCGAGGGTGCTTTTTGTGTTTCGGAGGTCCCTTCCTCATGGACGAGAGCTCCC  
GCCTGGCCTTTCCAGAGGCCACAGAGGAGGAGAAGCAGCTAAGCCTCACCATGATGGCCC  
AGTGGACCCACTTTGCCCGGACAGGGGACCCCAATAGCAAGGCTCTGCCTCCTTGGCCCC  
AATTCAACCAGGCGGAACAATATCTGGAGATCAACCCAGTGCCACGGGCCGGACAGAAGT  
TCAGGGAGGCCTGGATGCAGTTCTGGTCAGAGACGCTCCCCAGCAAGATACAACAGTGGC  
ACCAGAAGCAGAAGAACAGGAAGGCCAGGAGGACCTCTGAGGGCCAGGCCTGAACCTTCT  
TGGCTGGGGCAAACCACTCTTCAAGTGGTGGCAGAGTCCCAGCACGGCAGCCCGCCTCTC  
CCCCTGCTGAGACTTTAATCTCCACCAGCCCTTAAAGTGTGCGCCGCTCTGTGACTGGAG  
TTATGCTCTTTTGAATGTACAAGGCCGCTCCACCTCTGGGGCATTGTACAAGTTCT  
TCCCTCTCCCTGAAGTGCCTTTCTGCTTTCTTCGTGGTAGGTTCTAGCACATTCCTCTA  
GCTTCCTGGAGGACTCACTCCCCAGGAAGCCTTCCCTGCCTTCTCTGGGCTGTGCGGCCC  
CGAGTCTGCGTCCATTAGAGCACAGTCCACCCGAGGCTAGCACCGTGTCTGTGTCTGTCT  
CCCCCTCAGAGGAGCTCTCTCAAAATGGGGATTAGCCTAACCCCACTCTGTCACCCCACAC  
CAGGATCGGGTGGGACCTGGAGCTAGGGGGTGTGTTGCTGAGTGAGTGAGTGAAACACAGA  
ATATGGGAATGGCAGCTGCTGAACCTGAACCCAGAGCCTTCAGGTGCCAAAGCCATACTC  
AGGCCCCCACCAGCATTGTCCACCCTGGCCAGAAGGGTGCATGCCAATGGCAGAGACCTG  
GGATGGGAGAAGTCTTGGGGCGCCAGGGGATCCAGCCTAGAGCAGACCTTAGCCCCCTGAC  
TAAGGCCTCAGACTAGGGCGGGAGGGGTCTCCTCCTCTCTGCTGCCAGTCTTGGCCCCCT  
GCACAAGACAACAGAATCCATCAGGGCCATGAGTGTCAACCAGACCTGACCCTACCAAT  
TCCAGCCCCCTGACCCTCAGGACGCTGGATGCCAGCTCCCAGCCCCAGTGGCGGGTCTCTC  
CTCCCTTCTTGGCTTGGGGAGACCAGTTTCTGGGGAGCTTCCAAGAGCACCCACCAAGAC  
ACAGCAGGACAGGCCAGGGGAGGGCATCTGGACCAGGGCATCCGTGCGGCTATTGTCA  
GAGAAAAGAAGAGACCCACCCACTCGGGCTGCAAAAGGTGAAAAGCACCAAGAGGTTTTTC

277/392

AGATGGAAGTGAGAGGTGACAGTGTGCTGGCAGCCCTCACAGCCCTCGCTTGCTCTCCCT  
GCCGCTCTGCCTGGGCTCCCACTTTGGCAGCACTTGAGGAGCCCTTCAACCCGCCGCTG  
CACTGTAGGAGCCCCCTTCTGGGCTGGCCAAGGCCGGAGCCAGCTCCCTCAGCTTGCGGG  
GAGGTGCGGAGGGAGAGGGGCGGGCAGGAACCGGGGCTGCGCGCAGCGCTTGCGGGCCAG  
AGTGAGTTCGGGTGGGCGTGGGCTCGGCGGGGCCCCACTCAGAGCAGCTGGCCGGCCCC  
AGGCAGTGAGGGCCTTAGCACCTGGGCCAGCAGCTGCTGTGCTCGATTTCTCGCTGGGCC  
TTAGCTGCCTCCCCGCGGGCAGGGCTCGGGACCTGCAGCCCTCCATGCCTGACCCCTCC  
CCCACCCCCCGTGGGCTCCTGTGCGGCCGGAGCCTCCCCAAGGAGCGCCGCCCTGCTC  
CACAGCGCCCACTCCCATCGACCACCCAAGGGCTGAGGAGTGCGGGTGACAGCGCGGGA  
CTGGCAGGCAGCTCCACCTGCTGCCCCAGTGCTGGATCCACTGGGTGAAGCCAGCTGGGC  
TCCTGAGTCTGGTGGGACTTGGAGAACCTTTATGTCTAGCTAAGGGATTGTAAATACAC  
CGATGGGCACTCTGTATCTAGCTCAAGGTTTGTAACACACCAATCAGCACCCCTGTGTCT  
AGCTCAGTGTTTGTGAATGCACCAATCCACACTCTGTATCTGGCTACTCTGGTGGGGACT  
TGGAGAACCTTTGTGTCCACACTCTGTATCTAGCTAATCTAGTGGGGATGTGGAGAACCT  
TTGTGTCTAGCTCAGGGATCGTAAACGCACCAATCAGCACCCCTGTCAAAACAGACCACTT  
GACTCTCTGTAAATGGACCAATCAGCAGGATGTGGGTGGGGCGAGACAAGAGAATAAAA  
GCAGGCTGCCTGAGCCAGCAGTGACAACCCCCCTCGGGTCCCCTCCCACGCCGTGGAAGC  
TTTGTTCCTTCGCTCTTTGCAATAAATCTTGCTACTGCCCCAAA

278/392

**FIGURE 266**

MERAVRVESGVLVGVVCLLLACPATATGPEVAQPEVDTTLGRVRGRQVGVKGTDRLVNVF  
LGIPFAQPPLGPDRFSAPHPAQPWEGVRDASTAPPMCLQDVESMNSSRFVLNGKQQIFSV  
SEDCLVLNVYSPAIEVPAGSGRPVMVWVHGGALITGAATSYDGSALAAYGDVVVVTVQYRL  
GVLGFFSTGDEHAPGNQGFLLDVVAALRWVQENIAPFGGDLNCVTVFGGSSAGGSIISGLVL  
SPVAAGLFHRAITQSGVITTPGIIDSHPWPLAQKIAN TLACSSSSPAEMVQCLQQKEGEE  
LVLSKKLKNTIYPLTVDGTVFPKSPKELLKEKPFHSPFLMGVNNHEFSWLI PRGWGLLD  
TMEQMSREDMLAISTPVLTSLDVPPPEMMPTVIDEYLGSN SDAQAKCQAFQEFMGDVFINV  
PTVSFSRYLRDSGSPVFFYEFQHRPSSFAKIKPAWVKADHGAEGAFVFGGPFLMDESSRL  
AFPEATEEEKQLSLTMMAQWTHFARTGDPNSKALPPWPQFNQAEQYLEINPVPRAGQKFR  
EAWMQFWSETLPSKIQQWHQKQKNRKAQEDL

**Important features of the protein:**

**Signal peptide:**

amino acids 1-27

**Transmembrane domain:**

amino acids 226-245

**N-glycosylation site:**

amino acids 105-109

**N-myristoylation sites:**

amino acids 10-16, 49-55, 62-68, 86-92, 150-156, 155-161,  
162-168, 217-223, 227-233, 228-234, 232-238, 262-268, 357-363,  
461-467

**Prokaryotic membrane lipoprotein lipid attachment site:**

amino acids 12-23

**Carboxylesterases type-B serine active site:**

amino acids 216-232

279/392

## FIGURE 267

TGTCGCCTGGCCCTCGCCATGCGAGACCCCGCGAGCGTCCCCTCCCCGCCCCGGCCCTCCTG  
CTTCTGCTGCTGCTACTGGGGGGCGCCACGGCCTCTTTCCTGAGGAGCCGCCGCCGCTT  
AGCGTGGCCCCCAGGGACTACCTGAACCACTATCCCGTGTTTGTGGGCAGCGGGCCCGGA  
CGCCTGACCCCCGAGAAAGGTGCTGACGACCTCAACATCCAGCGAGTCTTGCGGGTCAAC  
AGGACGCTGTTTATTGGGGACAGGGACAACCTTACC CGTAGAGCTGGAGCCCCCACG  
TCCACGGAGCTGCGGTACCAGAGGAAGCTGACCTGGAGATCTAACCCAGCGACATAAAC  
GTGTGTGCGATGAAGGGCAAACAGGAGGGCGAGTGTGAAACTTCGTAAAGGTGCTGCTC  
CTTCGGGACGAGTCCACGCTCTTTGTGTGCGGTTCCAACGCCTTCAACCCGGTGTGCGCC  
AACTACAGCATAGACACCCTGCAGCCCGTCGGAGACAACATCAGCGGTATGGCCCGCTGC  
CCGTACGACCCCAAGCAGCCCAATGTTGCCCTCTTCTCTGACGGGATGCTCTTACAGCT  
ACTGTTACCGACTTCTAGCCATTGATGCTGTCTATACCAGCGCCTCGGGGACAGGCCC  
ACCCTGCGCACCGTGAAACATGACTCCAAGTGGTTCAAAGAGCCTTACTTTGTCCATGCG  
GTGGAGTGGGGCAGCCATGTCTACTTCTTCTTCCGGGAGATTGCGATGGAGTTAACTAC  
CTGGAGAAGGTGGTGGTGTCCCGCTGGCCCGAGTGTGCAAGAACGACGTGGGAGGCTCC  
CCCCGCGTGCTGGAGAAGCAGTGGACGTCCTTCTGAAGGCGCGGCTCAACTGCTCTGTA  
CCCGGAGACTCCCATTTCTACTTCAACGTGCTGCAGGCTGTCACGGGCGTGGTCAGCCTC  
GGGGGCCCGCCCGTGGTCTTGCCGTTTTCACGCCCAGCAACAGCATCCCTGGCTCG  
GCTGTCTGCGCCTTTGACCTGACACAGGTGGCAGCTGTGTTTGAAGGCCGCTTCCGAGAG  
CAGAAGTCCCCGAGTCCATCTGGACGCCGGTGCCGGAGGATCAGGTGCCTCGACCCCGG  
CCCGGGTGCTGCGCAGCCCCCGGGATGCAGTACAATGCCTCCAGCGCCTTGCCGGATGAC  
ATCCTCAACTTTGTCAAGACCCACCCTCTGATGGACGAGGCGGTGCCCTCGCTGGGCCAT  
GCGCCCTGGATCCTGCGGACCCTGATGAGGCACCAGCTGACTCGAGTGGCTGTGGACGTG  
GGAGCCGGCCCCCTGGGGCAACCAGACCCTTGTCTTCTGGGTTCTGAGGCGGGGACGGTC  
CTCAAGTTCCTCGTCCGGCCCCAATGCCAGCACCTCAGGGACGTCTGGGCTCAGTGTCTTC  
CTGGAGGAGTTTGAGACCTACCGGCCGGACAGGTGTGGACGGCCCCGGCGGTGGCGAGACA  
GGGCAGCGGCTGTGAGCTTGAGCTGGACGAGCTTCGGGGGGCCTGCTGGCTGCCTTC  
CCCCGCTGCGTGGTCCGAGTGCTGTGGCTCGCTGCCAGCAGTACTCGGGGTGATGAAG  
AACTGTATCGGCAGTCAGGACCCCTACTGCGGGTGGGCCCCCGACGCGCTCCTGCATCTTC  
CTCAGCCCGGGCACCAGAGCCGCTTTGAGCAGGACGTGTCCGGGGCCAGCACCTCAGGC  
TTAGGGGACTGCACAGGACTCCTGCGGGCCAGCCTCTCCGAGGACCGCGCGGGGCTGGTG  
TCGGTGAACCTGCTGTTAACGTGTCGGTGGCGGCCTTCGTGGTGGGAGCCGTGGTGTCC  
GGCTTCAGCGTGGGCTGGTTCGTGGGCCCTCCGTGAGCGGCGGGAGCTGGCCCGGCGCAAG  
GACAAGGAGGCCATCCTGGCGCACGGGGCGGGCGAGGCGGTGCTGAGCGTCAGCCGCTG  
GGCGAGCGCAGGGCGCAGGGTCCCGGGGGCGGGGCGGAGGCGGTGGCGGTGGCGCCGGG  
GTTCCCCCGGAGGCCCTGCTGGCGCCCCCTGATGCAGAACGGCTGGGCCAAGGCCACGCTG  
CTGCAGGGCGGGCCCCACGACCTGGACTCGGGGCTGCTGCCACGCCCGAGCAGACGCCG  
CTGCCGAGAAAGCGCCTGCCCCTCCGCACCCGCACCCCCACGCCCTGGGCCCCCGCGCC  
TGGGACCACGGCCACCCCTGCTCCCGCCTCCGCTTCATCCTCCCTCCTGCTGCTGGCG  
CCCGCCCGGGCCCCCGAGCAGCCCCCGCGCCTGGGGAGCCGACCCCCGAGCGCCGCTC  
TATGCTGCCCCGGCCCCGGCGCGCTCCACGGCGACTTCCCGCTACCCCCACGCCAGC  
CCGGACCGCCGGCGGGTGGTGTCCGCGCCACGGGCCCCCTTGACCCAGCCTCAGCCGCC  
GATGGCCTCCCGCGGCCCTGGAGCCCCCGCCCGACGGGCAGCCTGAGGAGGCCACTGGGC  
CCCCACGCCCTCCGGCCGCCACCCTGCGCCGCACCCACACGTTCAACAGCGGCGAGGCC  
CGGCCTGGGGACCGCCACCGCGGCTGCCACGCCCCGGCGGGCACAGACTTGGCCACCTC  
CTCCCCATATGGGGGGGCGGACAGGACTGCGCCCCCGTGCCCTAGGCCGGGGGCCCCCG  
ATGCCTTGGCAGTGCCAGCCACGGGAACAGGAGCGAGAGACGGTGCCAGAACGCCGGGG  
CCCCGGGCAACTCCGAGTGGGTGCTCAAGTCCCCCCCCGCGACCCACCCGCGAGTGGGGG

280/392

GCCCCCTCCGCCACAAGGAAGCACAAACCAGCTCGCCCTCCCCCTACCCGGGGCCGCAGGA  
CGCTGAGACGGTTTGGGGGTGGGTGGGCGGGAGGACTTTGCTATGGATTTGAGGTGACC  
TTATGCGCGTAGGTTTTGGTTTTTTTTTGCAGTTTTGGTTTCTTTTGCGGTTTTCTAACC  
AATTGCACAACCTCCGTTCTCGGGGTGGCGGCAGGCAGGGGAGGCTTGGACGCCGGTGGGG  
AATGGGGGGCCACAGCTGCAGACCTAAGCCCTCCCCCACCCTGGAAAGGTCCCTCCCCA  
ACCCAGGCCCCCTGGCGTGTGTGGGTGTGCGTGCGTGTGCGTGCCGTGTTTCGTGTGCAAGG  
GGCCGGGGAGGTGGGCGTGTGTGTGCGTGCCAGCGAAGGCTGCTGTGGGCGTGTGTGTCA  
AGTGGGCCACGCGTGCAGGGTGTGTGTCCACGAGCGACGATCGTGGTGGCCCCAGCGGCC  
TGGGCGTTGGCTGAGCCGACGCTGGGGCTTCCAGAAGGCCCGGGGGTCTCCGAGGTGCCG  
GTTAGGAGTTTGAACCCCCCCCCACTCTGCAGAGGGAAGCGGGGACAATGCCGGGGTTTCA  
GGCAGGAGACACGAGGAGGGCCTGCCCCGAAGTCACATCGGCAGCAGCTGTCTAAAGGGC  
TTGGGGGCCTGGGGGGCGGCGAAGGTGGGTGGGGCCCCCTCTGTAAATACGGCCCCAGGGT  
GGTGAGAGAGTCCCATGCCACCCGTCCCCTTGTGACCTCCCCCTATGACCTCCAGCTGA  
CCATGCATGCCACGTGGCTGGCTGGGTCTCTGCCCTCTTTGGAGTTTGCCTCCCCCAGC  
CCCCCTCCCCATCAATAAACTCTGTTTACAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA



281/392

**FIGURE 268**

MQTPRASPPRPALLLLLLLLLLGGAHGLFPPEPPPLSVAPRDYLNHYPVFGSGPGRLTPAE  
GADDLNIQVRVLRVNRTLFIGDRDNLRYVELEPPPTSTELRYQRKLTWRSNPSDINVCRMKG  
KQEGECRNFKVLLLLRDESTLFVCGSNAFNPVCANYSIDTLQPVGDNISGMARCPYDPKH  
ANVALFSDGMLFTATVTDFLAIDAVIYRSLGDRPTLRTVKHDSKWFKEPYFVHAVEWGS  
VYFFFREIAMEFNYLEKVVVSRVARVCKNDVGGSPRVLEKQWTSFLKARLNCSVPGDSHF  
YFNVLQAVTGVVSLGGRPVLAVFSTPSNSIPGSAVCAFDLTQVAAVFEGRFREQKSPES  
IWTPVPEDQVPRPRPGCCAAPGMQYNASSALPDDILNFVKTHPLMDEAVPSLGHPWILR  
TLMRHQLTRVAVDVGAGPWGNQTVVFLGSEAGTVLKFLVRPNASTSGTSGLSVFLEEFET  
YRPDRCGRPGGGETGQRLLSLELDAASGGLLAFFPRCVVRVPVARCQQYSGCMKNCIGSQ  
DPYCGWAPDGCIFLSPGTRAAFEQDVSGASTSGLGDCGTGLLRASLSEDRAGLVSVNLLV  
TSSVAAFVVGAVVSGFSVGVWFVGLRERRELARRKDKEAILAHGAGEAVLSVSRLGERRAQ  
GPGGRGGGGGGGAGVPPEALLAPLMQNGWAKATLLQGGPHDLDSGLLPTPEQTPLPQKRL  
PTPHPHPHALGPRAWDHGHPLLPASASSSLLLLAPARAPEQPPAPGEPTPDGRLYAARPG  
RASHGDFPLTPHASPDRRRVVSAPTGPLDPASAADGLPRPWSPPTGSLRRPLGPHAPPA  
ATLRRTHTFNSGEARPGDRHRGCHARPGTDLAHLPLPYGGADRTAPPVP

**Important features of the protein:**

**Signal peptide:**

amino acids 1-25

**Transmembrane domains:**

amino acids 318-339, 598-617

**N-glycosylation sites.**

amino acids 74-78, 155-159, 167-171, 291-295, 386-390,  
441-445, 462-466

**Glycosaminoglycan attachment sites:**

amino acids 51-55, 573-577

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

amino acids 102-106

**N-myristoylation sites:**

amino acids 21-27, 50-56, 189-195, 333-339, 382-388, 448-454,  
490-496, 491-497, 508-514, 509-515, 531-537, 558-564, 569-575,  
574-580, 580-586, 610-616, 643-649, 663-669, 666-672, 667-673,  
668-674, 669-675, 670-676, 868-874, 879-885

282/392

**FIGURE 269**

ATCTGAGTGAGCTAACTGACACAATGAAACTGTCAGGCATGTTTCTGCTCCTCTCTCTGG  
CTCTTTTCTGCTTTTAAACAGGTGTCTTCAGTCAGGGAGGACAGGTTGACTGTGGTGAGT  
TCCAGGACCCCAAGGTCTACTGCACTCGGGAATCTAACCCACACTGTGGCTCTGATGGCC  
AGACATATGGCAATAAATGTGCCTTCTGTAAGGCCATAGTGAAAAGTGGTGGAAAGATTA  
GCCTAAAGCATCCTGGAAAATGCTGAGTTAAAGCCAATGTTTCTTGGTGACTTGCCAGCT  
TTTGCAGCCTTCTTTTCTCACTTCTGCTTATACTTTTGCTGGTGGATTCCTTTAATTCAT  
AAAGACATACCTACTCTGCCTGGGTCTTGAGGAGTTCAATGTATGTCTATTTCTCTTGAT  
TCACTTGTCAATAAAGTACATTCTGCAAAAGCAAAA

283/392

## FIGURE 270

MKLSGMFLLLSLALFCFLTGVFSQGGQVDCGEFQDPKVYCTRESNPHCGSDGQTYGNKCA  
FCKAIVKSGGKISLKHPGKC

Important features of the protein:

Signal peptide:

amino acids 1-23

N-myristoylation sites:

amino acids 26-32, 52-58, 56-62, 69-75

Kazal serine protease inhibitors family signature:

amino acids 40-63

284/392

**FIGURE 271**

AACTTCTACATGGGCCTCCTGCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTAC  
ACCATCATGTCCCTCCCACCCTCCTTTGACTGCGGGCCGTTTCAGGTGCAGAGTCTCAGTT  
GCCCCGGAGCACCTCCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCCAGAATTCCA  
GTTCTGGTTTCATGCCAGCCTGTAAAAGGCCATGGAACTTTGGGTGAATCACCGATGCCA  
TTTAAGAGGGTCTTCTGCCAGGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTTCAT  
TTCAGTAGCCACCAGCCACCTGTGGCCGTTGAGTGCTTGAAATGAGGAACTGAGAAAATT  
AATTTCTCATGTATTTTCTCATTATTTATTAATTTTAACTGATAGTTGTACATATTT  
GGGGGTACATGTGATATTTGGATACATGTATACAATATATAATGATCAAATCAGGGTAAC  
TGGGATATCCATCACATCAAACATTTATTTTTTATTCTTTTATAGACAGAGTCTCACTCTG  
TCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAACCTCTGCCTGCCAGGT  
TCAAGCGATTCTCATGCCTCCACCTCCCAAGTAGCTGGGACTACAGGCATGCACCACAAT  
GCCCAACTAATTTTTGTATTTTTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGGCTGGC  
CTTGAACTCCTGGCCTCAAACAATCCACTTGCCTCGGCCTCCCAAAGTGTTATGATTACA  
GGCGTGAGCCACCGTGCCTGGCCTAAACATTTATCTTTTCTTTGTGTTGGGAACTTTGAA  
ATTATACAATGAATTATTGTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGGACT  
TCTTCCCTCTATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTC  
CTCTCTATCCTTCCCAACCTCTGATCACCTCATCTACTCTCTACCTCCATGAGATCCAC  
TTTTTTTAGCTCCCATGTGAGTAAGAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTA  
TTTCACTTAACATAATGACTTCCTGTTCCATCCATGTTGCTGCAAATGACAGGATTTTCGT  
TCTTAATTTCAATTAAAATAACCACACATGGCAAAAA

285/392

## FIGURE 272

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLV  
SCQPVKGHGTLGESPMFVKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site:

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature:

amino acids 68-79

286/392

**FIGURE 273**

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCG  
CAGCCTGGAGCATCTTCCTCATCGGGACTAAAATTGGGCTGTTCCCTTCAAGTAGCACCTC  
TATCAGTTATGGCTAAATCCTGTCCATCTGTGTGTCGCTGCGATGCGGGTTTCATTTACT  
GTAATGATCGCTTTCTGACATCCATTCCAACAGGAATACCAGAGGATGCTACAACCTCTCT  
ACCTTCAGAACAACCAAATAAATAATGCTGGGATTCCCTTCAGATTTGAAAACTTGCTGA  
AAGTAGAAAGAATATACCTATAACCACAACAGTTTAGATGAATTTCCCTACCAACCTCCCAA  
AGTATGTAAAAGAGTTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTAC  
TTTCAAAAATTCCTTATCTGGAAGAATTACATTTAGATGACAACTCTGTCTCTGCAGTTA  
GCATAGAAGAGGGAGCATTCGAGACAGCAACTATCTCCGACTGCTTTTCCTGTCCCGTA  
ATCACCTTAGCACAAATTCCTGCGGTTTGCCAGGACTATAGAAGAACTACGTTGGATG  
ATAATCGCATATCCACTATTTTCATCACCATCTCTTCAAGGTCTCACTAGTCTAAAACGCC  
TGGTTCTAGATGGAAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTTTCTTCAACC  
TAGTTAATTTGACAGAGCTGTCCCTGGTGCGGAATTCCCTGACTGCTGCACCAGTAAACC  
TTCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTGCCCC  
CAAATGCTTTTTCTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAA  
GTAATTTACCTCAGGGTATCTTTGATGATTTGGACAATATAACACAACCTGATTCTTCGCA  
ACAATCCCTGGTATTGCGGGTGCAAGATGAAATGGGTACGTGACTGGTTACAATCACTAC  
CTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAAAGGTTTCGTGGGATGG  
CTATTAAGGATCTCAATGCAGAACTGTTTGATTGTAAGGACAGTGGGATTGTAAGCACC  
TTCAGATAACCACTGCAATACCCAACACAGTGTATCCTGCCCAAGGACAGTGGCCAGCTC  
CAGTGACCAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAG  
GGAGTCCCTCAAGAAAAACAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATT  
ATATCTCTTGGAACCTTGCTCTACCTATGACTGCTTTGAGACTCAGCTGGCTTAAACTGG  
GCCATAGCCCGGCATTTGGATCTATAACAGAAACAATTGTAACAGGGGAACGCAGTGAGT  
ACTTGGTCACAGCCCTGGAGCCTGATTCACCCCTATAAAGTATGCATGGTTCCCATGGAAA  
CCAGCAACCTCTACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAACTGCACCCC  
TTCGAATGTACAACCCTACAACCACCCCTCAATCGAGAGCAAGAGAAAGAACCTTACAAA  
ACCCCAATTTACCTTTGGCTGCCATCATTTGGTGGGGCTGTGGCCCTGGTTACCATTGCC  
TTCTTGCTTTAGTGTGTGTTGGTATGTTTCATAGGAATGGATCGCTCTTCTCAAGGAACGTG  
CATATAGCAAAGGGAGGAGAAGAAAGGATGACTATGCAGAAGCTGGCACTAAGAAGGACA  
ACTCTATCCTGGAAATCAGGGAAACTTCTTTTCAGATGTTACCAATAAGCAATGAACCCA  
TCTCGAAGGAGGAGTTTGTAAATACACACCATATTTCTCCTAATGGAATGAATCTGTACA  
AAAACAATCACAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACT  
CAGATCACTCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTTTGGGTTTTTT  
AAACCTAAGGGAGGTGATGGT

287/392

**FIGURE 274**

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCND RFLTSIPTGIPEDA  
TTLYLQNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNI RTIT  
YDSLSKIPYLEELHDDNSVSAVSIIEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEEL  
RLDDNRISTISSPSLQGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAA  
PVNLPGTNLRKLYLQDNHINRVPPNAFSYLRLQLYRLDMSNNLSNLPQGIFDDLNDITQL  
ILRNNPWYCGCKMKWVRDWLQSLPVKVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGI  
VSTIQITTAIPNTVYPAQGQWPAPVTKQPDIKNPKLTKDQQTGSPSRKTITITVKS VTS  
DTIHISWKLALPMTALRLSWLKLGHSPAFGSITETIVTGERSEYLVTALEPDSPYKVC MV  
PMETS NLYLFDETPVC IETETAPLRMYNPTTTLNREQEKEPYKNPNLPLAAIIGGAVALV  
TIAL LALVCWYVHRNGSLFSRNCAYS KGRRRKDDYAEAGTKKDNSILEIRETSFQMLPIS  
NEPISKEEFVIHTIFPPNGMNL YKNNHSESSSNRSYRDSGIPDS DSHSHS

**Important features of the protein:****Signal peptide:**

amino acids 1-28

**Transmembrane domain:**

amino acids 531-552

**N-glycosylation sites:**

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

**Tyrosine kinase phosphorylation site:**

amino acids 515-522

**N-myristoylation sites:**

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561, 640-645

**Amidation site:**

amino acids 567-570

**Leucine zipper pattern:**

amino acids 159-180

**Phospholipase A2 aspartic acid active site:**

amino acids 34-44

288/392

**FIGURE 275**

AGGGCCCCGCGGGTGGAGAGAGCGACGCCCCGAGGGGATGGCGGCAGCGTCCCGGAGCGCCT  
CTGGCTGGGCGCTACTGCTGCTGGTGGCACTTTGGCAGCAGCGCGCGGCCGCTCCGGCG  
TCTTCCAGCTGCAGCTGCAGGAGTTTCATCAACGAGCGCGGCGTACTGGCCAGTGGGCGGC  
CTTGCGAGCCCCGGCTGCCGGACTTTCTTCCGCGTCTGCCTTAAGCACTTCCAGGCGGTG  
TCTCGCCCGGACCCTGCACCTTCGGGACCGTCTCCACGCCGGTATTGGGCACCAACTCCT  
TCGCTGTCCGGGACGACAGTAGCGGCGGGGGCGCAACCTCTCCAAGTGCCTTCAATT  
TCACCTGGCCGGGTACCTTCTCGCTCATCATCGAAGCTTGGCACGCGCCAGGAGACGACC  
TGCGGCCAGAGGCCTTGCCACCAGATGCACTCATCAGCAAGATCGCCATCCAGGGCTCCC  
TAGCTGTGGGTGAGAACTGGTTATTGGATGAGCAAACCAGCACCTCACAAGGCTGCGCT  
ACTCTTACCGGGTCATCTGCAGTGACAATACTATGGAGACAAGTCTCCGCGCTGTGCA  
AGAAGCCCAATGACCCTTCGGCCACTATGTGTGCCAGCCAGATGGCAACTGTCTGCTGCC  
TGCCCCGGTTGGACTGGGGAATATTGCCAACAGCCCTATCTGTCTTTTCGGGCTGTCTGAAC  
AGAATGGCTACTGCAGCAAGCCAGCAGAGTGCCTCTGCCGCCAGGCTGGCAGGGCCGGC  
TGTGTAACGAATGCATCCCCCACAATGGCTGTGCCACGGCACCTGCAGCACTCCCTGGC  
AATGTACTTGTGATGAGGGCTGGGGAGGCCTGTTTTGTGACCAAGATCTCAACTACTGCA  
CCCACCACTCCCCATGCAAGAATGGGGCAACGTGCTCCAACAGTGGGCAGCGAAGCTACA  
CCTGCACCTGTGCCCGAGGTACACTGGTGTGGACTGTGAGCTGGAGCTCAGCGAGTGTG  
ACAGCAACCCCTGTGCAATGGAGGCAGCTGTAAGGACCAGGAGGATGGCTACCACTGCC  
TGTGTCTCCGGGCTACTATGGCCTGCAGTGTGAACACAGCACCTTGAGCTGCGCCGACT  
CCCCCTGCTTCAATGGGGGCTCCTGCCGGGAGCGCAACCAGGGGGCCAAGTATGCTTGTG  
AATGTCCCCCAACTTCACCGGCTCCAAGTGCAGAGAAGAAAGTGGACAGGTGCACCAGCA  
ACCCCTGTGCCAACGGGGACAGTGCCTGAACCGAGGTCCAAGCCGCATGTGCCGCTGCC  
GTCCTGGATTTCACGGGCACCTACTGTGAAGTCCACGTCAGCGACTGTGCCCCGTAACCTT  
GCGCCACGGTGGCACTTGCCATGACCTGGAGAATGGGCTCATGTGCACCTGCCCTGCCG  
GCTTCTCTGGCCGACGCTGTGAGGTGCGGACATCCATCGATGCCTGTGCCTCGAGTCCCT  
GCTTCAACAGGGCCACCTGCTACACCGACCTCTCCACAGACACCTTTGTGTGCAACTGCC  
CTTATGGCTTTGTGGGCAGCCGCTGCGAGTTCCCCGTGGGCTTGCCGCCCAGCTTCCCCT  
GGGTGGCCGTCTCGCTGGGTGTGGGGCTGGCAGTGTGCTGGTACTGCTGGGCATGGTGG  
CAGTGGCTGTGCGGCAGCTGCGGCTTCGACGGCCGGACGACGGCAGCAGGGAAGCCATGA  
ACAACTTGTGCGACTTCCAGAAGGACAACCTGATTCCTGCGCCAGCTTAAAAACACAA  
ACCAGAAGAAGGAGCTGGAAGTGGACTGTGGCCTGGACAAGTCCAAGTGTGGCAACAGC  
AAAACCACACATTGGACTATAATCTGGCCCCAGGGCCCCCTGGGGCGGGGACCATGCCAG  
GAAAGTTTCCCCACAGTGACAAGAGCTTAGGAGAGAAGGCGCCACTGCGGTTACACAGTG  
AAAAGCCAGAGTGTGCGATATCAGCGATATGCTCCCCAGGGACTCCATGTACCAGTCTG  
TGTGTTTGATATCAGAGGAGAGGAATGAATGTGTGCTTGGCACGGAGGTATAAGGCAGGA  
GCCTACCTGGACATCCCTGCTCAGCCCCGCGGCTGGACCTTCCTTCTGCATTGTTTACA



289/392

**FIGURE 276**

MAAASRSASGWALLLLVALWQORAAGSGVFQLQLQEFINERGVLASGRPCEPGCARTFFRV  
CLKHFQAVVSPGPCTFGTVSTPVLGTNSFAVRDSSGGGRNPLQLPFNFTWPGTFSLIIE  
AWHAPGDDLRLPEALPPDALISKIAIQGSLAVGQNWLLDEQTSTLTRLRYSYRVICSDNY  
GDNC SRLCKKRNDHFHGYVCQPDGNLSCLPGWTGEYCQQPICLSGCHEQNGYCSKPAECL  
CRPGWQGRLCNECIPHNGCRHGTCTSPWQCTCDEGWGGLFCDQDLNYCTHHSPCKNGATC  
SNSGQRSYTCTCRPGYTGVDCELELSECDSPNCRNGGSKDQEDGYHCLCPPGYGLHCE  
HSTLSCADSPCFNGGSCRERNQGANACECPNFTGSNCEKKVDRCTSNPCANGGQCLNR  
GPSRMCRCRPGFTGTCELVSDCARNPCAHGGTCHDLENGLMCTCPAGFSGRRCEVRTS  
IDACASSPCFNRTATCYTDLSTDTFVCNCPYGFVGSRCFVPVGLPPSFPWVAVSLGVGLAV  
LLVLLGMVAVAVRQLRLRRPDDGSREAMNNLSDFQKDNLI PAAQLKNTNQKKELEVDCGL  
DKSNC GKQQNHTLDYNLAPGPLGRGTMPGKFPHSDKSLGEKAPLRLHSEKPECRISAICS  
PRDSMYQSVCLISEERNECVIATEV

**Important features of the protein:****Signal peptide:**

amino acids 1-26

**Transmembrane domain:**

amino acids 530-552

**N-glycosylation sites:**amino acids 108-112, 183-187, 205-209, 393-397, 570-574,  
610-614**Glycosaminoglycan attachment site:**

amino acids 96-100

**Tyrosine kinase phosphorylation site:**

amino acids 340-347

**N-myristoylation sites:**amino acids 42-48, 204-210, 258-264, 277-283, 297-303,  
383-389, 415-421, 461-467, 522-528, 535-541, 563-569,  
599-605, 625-631**Amidation site:**

amino acids 471-475

**Aspartic acid and asparagine hydroxylation site:**

amino acids 339-351

**EGF-like domain cysteine pattern signature:**amino acids 173-185, 206-218, 239-251, 270-282, 310-322,  
348-360, 388-400, 426-438, 464-476, 506-518**Calcium-binding EGF-like:**

amino acids 224-245, 255-276, 295-316, 333-354, 373-394,

290/392

**FIGURE 277**

GGCAGTGCAGCCGCCTCACAGGTGCGCGGACGGGCCAGGCGGGCGGCCTCCTGAACCGAA  
CCGAATCGGCTCCTCGGGCCGTGTCCTCCCGCCCCCTCCTCGCCCGCCGCGGAGTTTTC  
TTTCGGTTTCTTCCAAGATTCTTGGCCTTCCCTCGACGGAGCCGGGCCCAGTGCGGGGGC  
GCAGGGCGCGGGAGCTCCACCTCCTCGGCTTTCCCTGCGTCCAGAGGCTGGCATGGCGCG  
GGCCGAGTACTGAGCGCACGGTCGGGGCACAGCAGGGCCGGGGGTGCAGCTGGCTCGCG  
CCTCCTCTCCGGCCGCCGTCTCCTCCGGTCCCTGGCGAAAGCCATTGAGACACCAGCTGG  
ACGTCACGCGCCGGAGCATGTCTGGGAGTCAGAGCGAGGTGGCTCCATCCCCGCAGAGTC  
CGCGGAGCCCCGAGATGGGACGGGACTTGCGGGCCGGGTCCCGCGTGCTCCTGCTCCTGC  
TTCTGCTCCTGCTGGTGTACCTGACTCAGCCAGGCAATGGCAACGAGGGCAGCGTCACTG  
GAAGTTGTTATTGTGGTAAAAGAATTTCTTCCGACTCCCCGCCATCGGTTCA GTTCATGA  
ATCGTCTCCGGAAACACCTGAGAGCTTACCATCGGTGTCTATACTACACGAGGTTCCAGC  
TCCTTTCTGAGCGTGTGTGGGGGCAACAAGGACCCATGGGTTCAAGGAATTGATGAGCT  
GTCTTGATCTCAAAGAATGTGGACATGCTTACTCGGGGATTGTGGCCCACCAGAAGCATT  
TACTTCCTACCAGCCCCCAATTTCTCAGGCCTCAGAGGGGGCATCTTCAGATATCCACA  
CCCCTGCCCAGATGCTCCTGTCCACCTTGCACTCAGCGCCCCACCCTCCAGTAG  
GATCACTGTCCTCGGACAAAGAGCTCACTCGTCCCAATGAAACCACCATTCACTGCGG  
GCCACAGTCTGGCAGCTGGGCCTGAGGCTGGGGAGAACCAGAAGCAGCCGGAAAAAATG  
CTGGTCCCACAGCCAGGACATCAGCCACAGTGCCAGTCCTGTGCCTCCTGGCCATCATCT  
TCATCCTCACCGCAGCCCTTTCTATGTGCTGTGCAAGAGGAGGAGGGGGCAGTCACCGC  
AGTCCTCTCCAGATCTGCCGGTTCATTATATACCTGTGGCACCTGACTCTAATACCTGAG  
CCAAGAATGGAAGCTTGTGAGGGTAAACTGTGGCTTATTCTTACAAAAGTGAATAAAG  
GAGACTGACCCCTGACAACATGGTAGGCACTGTAAAAA

291/392

**FIGURE 278**

MGRDLRPGSRVLLLLLLLLLVYLTQPGNGNEGSVTGSCYCGKRISSDSPPSVQFMNRLRK  
HLRAYHRCLYYTRFQLLSWSVCGGNKDPWVQELMSCLDLKECGHAYSGIVAHQKHLPTS  
PPISQASEGASSDIHTPAQMLLSTLQSTQRPTLPVGSLSDDKELTRPNETTIHTAGHSLA  
AGPEAGENQKQPEKNAGPTARTSATVPVLCLLAIFILTAALSYVLCKRRRGQSPQSSPD  
LPVHYIPVAPDSNT

**Important features of the protein:**

**Signal peptide:**

1-26

**Transmembrane domain:**

204-223

**N-glycosylation site:**

168-172

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

42-46

**N-myristoylation site:**

29-35, 32-38, 36-42, 156-162

**Amidation site:**

40-44

292/392

**FIGURE 279**

CGCGAGGCGCGGGGAGCCTGGGACCAGGAGCGAGAGCCGCTACCTGCAGCCGCCGCCCA  
CGGCACGGCAGCCACCATGGCGCTCCTGCTGTGCTTCGTGCTCCTGTGCGGAGTAGTGGA  
TTTCGCCAGAAGTTTGAGTATCACTACTCCTGAAGAGATGATTGAAAAAGCCAAAGGGGA  
AACTGCCTATCTGCCATGCAAATTTACGCTTAGTCCCGAAGACCAGGGACCGCTGGACAT  
CGAGTGGCTGATATCACCAGCTGATAATCAGAAGGTGGATCAAGTGATTATTTTATATTC  
TGGAGACAAAATTTATGATGACTACTATCCAGATCTGAAAGGCCGAGTACATTTTACGAG  
TAATGATCTCAAATCTGGTGATGCATCAATAAATGTAACGAATTTACAACGTGCAGATAT  
TGGCACATATCAGTGCAAAGTGAAAAAGCTCCTGGTGTTGCAAATAAGAAGATTTCATCT  
GGTAGTTCTTGTTAAGCCTTCAGGTGCGAGATGTTACGTTGATGGATCTGAAGAAATTGG  
AAGTGACTTTAAGATAAAATGTGAACCAAAAGAAGGTTCACTTCCATTACAGTATGAGTG  
GCAAAAATTTGCTGACTCACAGAAAATGCCACTTCATGGTTAGCAGAAATGACTTCATC  
TGTTATATCTGTAAAAAATGCCTCTTCTGAGTACTCTGGGACATACAGCTGTACAGTCAG  
AAACAGAGTGGGCTCTGATCAGTGCCTGTTGCGTCTAAACGTTGTCCCTCCTTCAAATAA  
AGCTGGACTAATTGCAGGAGCCATTATAGGAACCTTTGCTTGCTCTAGCGCTCATTGGTCT  
TATCATCTTTTGCTGTCGTAAAAAGCGCAGAGAAGAAAAATATGAAAAGGAAGTTCATCA  
CGATATCAGGGAAGATGTGCCACCTCCAAAGAGCCGTACGTCCACTGCCAGAAGCTACAT  
CGGCAGTAATCATTCATCCCTGGGGTCCATGTCTCCTTCCAACATGGAAGGATATTCCAA  
GACTCAGTATAACCAAGTACCAAGTGAAGACTTTGAACGCACCTCCTCAGAGTCCGACTCT  
CCCACCTGCTAAGTTCAAGTACCCTTACAAGACTGATGGAATTACAGTTGTATAAATATG  
GACTACTGAAGAATCTGAAGTATTGTATTATTTGACTTTATTTTAGGCCCTCTAGTAAAGA  
CTTAAATGTTTTTTAAAAAAGCACAGGCACAGAGATTAGAGCAGCTGTAAGAACACAT  
CTACTTTATGCAATGGCATTAGACATGTAAGTCAGATGTCATGTCAAAATTAGTACGAGC  
CAAATTCCTTTGTTAAAAAACCTATGTATAGTGACACTGATAGTTAAAGATGTTTTATT  
ATATTTTCAATAACTACCCTAACAAATTTTAACTTTTCATATGCATATCTGATATGT  
GGTCTTTTAGGAAAAGTATGGTTAATAGTTGATTTTTTCAAAGGAAATTTTAAATTCCTTA  
CGTTCTGTTTAAATGTTTTTGCTATTTAGTTAAATACATTGAAGGGAATACCCGTTCTTT  
TCCCTTTTATGCACACAACAGAAACACGCGTTGTGCATGCCTCAAACATTTTTTTATTTG  
CAACTACATGATTTTACACAATTCTCTTAAACAACGACATAAAATAGATTTCTTGTATA  
TAAATAACTTACATACGCTCCATAAAGTAAATTTCTCAAAGGTGCTAGAACAAATCGTCCA  
CTTCTACAGTGTTCTCGTATCCAACAGAGTTGATGCACAATATATAAATACTCAAGTCCA  
ATATTAAAACTTAGGCCTTGACTAACTTTAATAAAATTTCTCAAACATATCAATATC  
TAAAGTGATATATTTTTTAAAGAAAGATTATTTCTCAATAACTTCTATAAAAAATAAGTTTG  
ATGGTTTGGCCCATCTAACTTCACTACTATTAGTAAGAACTTTTAACTTTTAAATGTGTAG  
TAAGGTTTATTCTACCTTTTTCTCAACATGACACCAACACAATCAAAAACGAAGTTAGTG  
AGGTGCTAACATGTGAGGATTAATCCAGTGATTCCGGTCACAATGCATTCCAGGAGGAGG  
TACCCATGTCACTGGAATTGGGCGATATGGTTTTATTTTTTCTTCCCTGATTTGGATAACC  
AAATGGAACAGGAGGAGGATAGTGATTCTGATGGCCATTCCCTCGATACATTCCCTGGCTT  
TTTTCTGGGCAAAGGGTGCCACATTGGAAGAGGTGGAAATATAAGTTCTGAAATCTGTAG  
GGAAGAGAACACATTAAGTTAATTCAAAGGAAAAAATCATCATCTATGTTCCAGATTTCT  
CATTAAAGACAAAGTTACCCACAACACTGAGATCACATCTAAGTGACACTCCTATTGTCA  
GGTCTAAATACATTAAAAACCTCATGTGTAATAGGCGTATAATGTATAACAGGTGACCAA  
TGTTTTCTGAATGCATAAAGAAATGAATAAACTCAAACACAGTACTTCCTAAACAACCTTC  
AACCAAAAAAGACCAAAACATGGAACGAATGGAAGCTTGTAAGGACATGCTTGTTTTAGT  
CCAGTGGTTTCCACAGCTGGCTAAGCCAGGAGTCACTTGAGGCTTTTAAATACAAAACA  
TTGGAGCTGGAGGCCATTATCCTTAGCAAACTAATGCAGAAACAGAAAATCAACTACCGC  
ATGTTCTCACTTATAAGTGGGAGGTAATGATAAGAACTTATGAACACAAGAAGGAAACA  
ATAGACATTGGAGTCTATTTGAGAGGGGAGGGTGGGAGAAGGAAAAGGAGCAGAAAAGAT  
AACTATTGAGTACTGCCTTCACACCTGGGTGATGAAATAATATGTACAACAAATCCCTGT  
GACACATGTTTACCTATGGAACAAACCTTCATGTGTATCCCTAAACCTAAAAATAAAGTT

293/392

**FIGURE 280**

MALLLCFVLLCGVVDFARSLSITTPPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLIS  
PADNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC  
KVKKAPGVANKKIHLVVLVKPSGARCIVDGSEEIGSDFKIKCEPKEGSLPQYEWQKLS  
SQKMPTSWLAEMTSSVISVKNASSEYSGTYSCTVRNRVGSQCLLRNLNVPPSNKAGLIA  
GAIIGTLLALALIGLIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHS  
SLGSMSPSNMEGYSKTQYNQVPSEDFERTPQSPTLPPAKFKYPYKTDGITVV

**Signal sequence.**

amino acids 1-19

**Transmembrane domain:**

amino acids 236-257

**N-glycosylation sites:**

amino acids 106-110, 201-205, 298-302

**Tyrosine kinase phosphorylation sites:**

amino acids 31-39, 78-85, 262-270

**N-myristoylation sites:**amino acids 116-122, 208-214, 219-225, 237-243, 241-247,  
245-251, 296-302**Myelin P0 protein:**

amino acids 96-125

294/392

**FIGURE 281**

TGCATCAGTGCCCAGGCAAGCCCAGGAGTTGACATTTCTCTGCCCAGCCATGGGCCTCAC  
CCTGCTCTTGCTGCTGCTCCTGGGACTAGAAGGTCAGGGCATAGTTGGCAGCCTCCCTGA  
GGTGCTGCAGGCACCCGTGGGAAGCTCCATTCTGGTGCACTGCCACTACAGGCTCCAGGA  
TGTCAAAGCTCAGAAGGTGTGGTGCCGGTTCTTGCCGGAGGGGTGCCAGCCCCCTGGTGTC  
CTCAGCTGTGGATCGCAGAGCTCCAGCGGGCAGGCGTACGTTTCTCACAGACCTGGGTGG  
GGGCCTGCTGCAGGTGGAAATGGTTACCCTGCAGGAAGAGGATGCTGGCGAGTATGGCTG  
CATGGTGGATGGGGCCAGGGGGCCCCAGATTTTGACAGAGTCTCTCTGAACATACTGCC  
CCCAGAGGAAGAAGAAGAGACCCATAAGATTGGCAGTCTGGCTGAGAACGCATTCTCAGA  
CCCTGCAGGCAGTGCCAACCCTTTGGAACCCAGCCAGGATGAGAAGAGCATCCCCCTTGAT  
CTGGGGTGCTGTGCTCCTGGTAGGTCTGCTGGTGGCAGCGGTGGTGCTGTTTGCTGTGAT  
GGCCAAGAGGAAACAAGAATCCCTCCTCAGTGGTCCACCACGTCAGTGACTCTGGACCGG  
CTGCTGAATTGCCTTTGGATGTACCACACATTAGGCTTGACTCACCACCTTCATTTGACA  
ATACCACCTACACCAGCCTACCTCTTGATTCCCCATCAGGAAAACCTTCACCTCCAGCTC  
CATCCTCATTGCCCCCTCTACCTCCTAAGGTCCTGGTCTGCTCCAAGCCTGTGACATATG  
CCACAGTAATCTTCCCGGGAGGGAACAAGGGTGGAGGGACCTCGTGTGGGCCAGCCCAGA  
ATCCACCTAACAATCAGACTCCATCCAGCTAAGCTGCTCATCACACTTTAAACTCATGAG  
GACCATCCCTAGGGGTCTGTGCATCCATCCAGCCAGCTCATGCCCTAGGATCCTTAGGA  
TATCTGAGCAACCAGGGACTTTAAGATCTAATCCAATGTCCTAACTTTACTAGGGAAAGT  
GACGCTCAGACATGACTGAGATGTCTTGGGGAAGACCTCCCTGCACCCAACTCCCCACT  
GGTTCTTCTACCATTACACACTGGGCTAAATAAACCTAATAATGATGTGCAAAAAAAAAA  
AA

295/392

## FIGURE 282

MGLTLLLLLLLLLGLGQGIVGSLPEVLQAPVGSSILVQCHYRLQDVKAQKVWCRFLPEGCQ  
PLVSSAVDRRAPAGRRTFLTDLGGGLLQVEMVTLQEEDAGEYGCMVDGARGPQILHRVSL  
NILPPEEEEEETHKIGSLAENAFSDPAGSANPLEPSQDEKSIPLIWGAVLLVGLLVAAVVL  
FAVMAKRKQESLLSGPPRQ

**Important features of the protein:**

**Signal peptide:**

amino acids 1-15

**Transmembrane domain:**

amino acids 161-181

**N-myristoylation sites:**

amino acids 17-23, 172-178

**Amidation site:**

amino acids 73-79

296/392

**FIGURE 283**

GTAGCATAGTGTGCAGTTCACCTGGACCAAAGCTTTGGCTGCACCTCTTCTGGAAAGCTG  
GCCATGGGGCTCTTCATGATCATTGCAATTCTGCTGTTCCAGAAACCCACAGTAACCGAA  
CAACTTAAGAAGTGCTGGAATAACTATGTACAAGGACATTGCAGGAAAATCTGCAGAGTA  
AATGAAGTGCCTGAGGCACTATGTGAAAATGGGAGATACTGTTGCCTCAATATCAAGGAA  
CTGGAAGCATGTAAAAAATTACAAAGCCACCTCGTCCAAAGCCAGCAACACTTGCACTG  
ACTCTTCAAGACTATGTTACAATAATAGAAAATTTCCAAGCCTGAAGACACAGTCTACA  
TAAATCAAATACAATTTTCGTTTTCACTTGCTTCTCAACCTAGTCTAATAAACTAAGGTGA  
TGAGATATACATCTTCTTCCTTCTGGTTTCTTGATCCTTAAAATGACCTTCGAGCATATT  
CTAATAAAGTGCATTGCCAGTTAAAAA



297/392

## FIGURE 284

MGLFMIIAILLFQKPTVTEQLKKCWNNYVQGHCRKICRVNEVPEALCENGRYCCLNIKEL  
EACKKITKPPRPKPATLALTLQDYVTIIEFNSLKTQST

Important features of the protein:

Signal peptide:

None

Transmembrane domain:

None

cAMP- and cGMP-dependent protein kinase phosphorylation site:

64-68

298/392

**FIGURE 285**

GATGGCGCAGCCACAGCTTCTGTGAGATTTCGATTTCTCCCCAGTTCCCCTGTGGGTCTGA  
GGGGACCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTC  
CCCAAAACAAGTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTG  
CCTGCTGTTCCAGGCCTTACCTGCTGGGCACCTAACGGCGGAGCCAGGATGGGGACAGAAT  
AAAGGAGCCACGACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATC  
TTCTCTTCACGGGAGGCTTGGCAGTTTTTCTTACTCCTGTGGTCTCCAGATTTCAAGGCCT  
AAGATGAAAGCCTCTAGTCTTGCCTTCAGCCTTCTCTCTGCTGCGTTTTATCTCCTATGG  
ACTCCTTCCACTGGACTGAAGACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTT  
CAGGAAATACGAAATGGATTTTTCTGAGATACGGGGCAGTGTGCAAGCCAAAGATGGAAAC  
ATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACAAAGCCTGCGAATCGA  
TGCTGCCTCCTGCGCCATTTGCTAAGACTCTATCTGGACAGGGTATTTAAAAACTACCAG  
ACCCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTCGCCAATTCCTTTCTTACCATC  
AAGAAGGACCTCCGGCTCTCTCATGCCCACATGACATGCCATTGTGGGGAGGAAGCAATG  
AAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTG  
AAGGCTTTGGGGGAAGTAGACATTCTTCTGCAATGGATGGAGGAGACAGAATAGGAGGAA  
AGTGATGCTGCTGCTAAGAATATTGAGGTCAGAGCTCCAGTCTTCAATACCTGCAGAG  
GAGGCATGACCCCAAACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCAAGTGTA  
TCTTATTTATGCATTACTTGCTTCCTTGCAATGATTTGTCTTTATGCATCCCCAATCTTAAT  
TGAGACCATACTTGTATAAGATTTTTTGTAATATCTTTCTGCTATTGGATATATTTATTAG  
TTAATATATTTATTTATTTTTTGCTATTTAATGTATTTATTTTTTTACTTGGACATGAAA  
CTTTAAAAAAATTACAGATTATATTTATAACCTGACTAGAGCAGGTGATGTATTTTTAT  
ACAGTAAAAAATAACCTTGTAATTTCTAGAAGAGTGGCTAGGGGGGTATTTCATTTG  
TATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGATATTTGAAATGAACC  
AATGACTACTTAGGATGGGTGTGGAATAAGTTTTGATGTGGAATTGCACATCTACCTTA  
CAATTACTGACCATCCCAGTAGACTCCCAGTCCCATAATTGTGTATCTTCCAGCCAGG  
AATCCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATACCAAAAAAA  
AAAAA

299/392

## FIGURE 286

MKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNGFSEIRGSVQAKDGN  
DIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKISSLANSTLIK  
KDLRLSHAHMTCHCGEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEETE

300/392

**FIGURE 287**

AATGCCCCATGCGCACCCACAGCTCGCGCTCCTGCAAGTGTTCTTTCTGGTGTTCCCCG  
ATGGCGTCCGGCCTCAGCCCTCTTCTCCCCATCAGGGGCAGTGCCACGTTCTTTGGAGC  
TGCAGCGAGGGACGGATGGCGGAACCTCCAGTCCCCTTCAGAGGCGACTGCAACTCGCC  
CGGCCGTGCCTGGACTCCCTACAGTGGTCCCTACTCTCGTGACTCCCTCGGCCCCCTGGGA  
ATAGGACTGTGGACCTCTTCCCAGTCTTACCGATCTGTGTCTGTGACTTGACTCCTGGAG  
CCTGCGATATAAATTGCTGCTGCGACAGGGACTGCTATCTTCTCCATCCGAGGACAGTTT  
TCTCCTTCTGCCTTCCAGGCAGCGTAAGGTCTTCAAGCTGGGTTTGTGTAGACAACTCTG  
TTATCTTCAGGAGTAATTCCCCGTTTCTTCAAGAGTTTTCATGGATTCTAATGGAATCA  
GGCAGTTTTGTGTCCATGTGAACAACCTCAAACCTTAACTATTTCCAGAAGCTTCAAAAGG  
TCAATGCAACCAACTTCCAGGCCCTGGCTGCAGAGTTTGGAGGCGAATCATTCACTTCAA  
CATTCCAAACTCAATCACCACCATCTTTTTTACAGGGCTGGGGACCCCATTTCTACTTACT  
TCCCCAAGTGGTCTGTAATAAGCTTGCTGAGACAACCTGCAGGAGTTGGAGCTGGGGGAC  
TCTGTGCTGAAAGCAATCCTGCAGGTTTCTTAGAGAGTAAAAGTACAACCTGCACTCGTT  
TTTTCAAGAACCTGGCTAGTAGCTGTACCTTGGATTGAGCCCTCAATGCTGCCTCTTACT  
ATAACTTCACAGTCTTAAAGGTTCCAAGAAGCATGACTGATCCACAGAATATGGAGTTCC  
AGGTTCTGTAAATACTTACCTCACAGGCTAATGCTCCTCTGTTGGCTGGAAACACTTGTC  
AGAATGTAGTTTCTCAGGTCACCTATGAGATAGAGACCAATGGGACTTTTGGAAATCCAGA  
AAGTTTCTGTGAGTTTGGGACAAACCAACCTGACTGTTGAGCCAGGCGCTTCCTTACAGC  
AACACTTCATCCTTCGCTTCAGGGCTTTTCAACAGAGCACAGCTGCTTCTCTCACCAGTC  
CTAGAAGTGGGAATCCTGGCTATATAGTTGGGAAGCCACTCTTGGCTCTGACTGATGATA  
TAAGTTACTCAATGACCCTCTTACAGAGCCAGGGTAATGGAAGTTGCTCTGTTAAAAGAC  
ATGAAGTGCAGTTTGGAGTGAATGCAATATCTGGATGCAAGCTCAGGTTGAAGAAGGCAG  
ACTGCAGCCACTTGCAGCAGGAGATTTATCAGACTCTTCATGGAAGGCCAGACCAGAGT  
ATGTTGCCATCTTTGGTAATGCTGACCCAGCCAGAAAGGAGGGTGGACCAGGATCCTCA  
ACAGGCACTGCAGCATTTAGCTATAAACTGTACTTCTGCTGTCTCATAACAGTTTCCC  
TGGAGATCCAGGTATTGTGGGCATATGTAGGTCTCCTGTCCAACCCGCAAGCTCATGTAT  
CAGGAGTTCGATTCTTATACCAGTGCCAGTCTATACAGGATTCTCAGCAAGTTACAGAAG  
TATCTTTGACAACTCTTGTGAACCTTGTGGACATTACCCAGAAGCCACAGCCTCCAAGGG  
GCCAACCCAAAATGGACTGGAAATGGCCATTGCACTTCTTCCCTTCAAAGTGGCATTCA  
GCAGAGGAGTATTCTCTCAAAAATGCTCAGTCTCTCCATCCTTATCCTGTGCCTCTTAC  
TACTTGGAGTTCTCAACCTAGAGACTATGTGAAGAAAAGAAAATAATCAGATTTCAGTTT  
TCCCTATGAGAACTCTGAGGCAGCCACTTATCTTGGCTAAATAGAACCTCACCTGCTCA  
TGACCAGAGAGCATTTAGGATAATAGATGACCTAACTGAAGGAATCCTTGTATATGAAAG  
GAGTTATTTAGAAAAGCAATAAAAAATATTTTATTCATCNTAAAAAAAAA

301/392

**FIGURE 288**

MRTPLQALLQVFFLVFPDGVPRQPSSSPSGAVPTSLELQRGTDGGTLQSPSEATATRPVAV  
PGLPTVVPTLVTPSAPGNRTVDLFPVLPICVCDLTPGACDINCCCDRDCYLLHPRTVFSF  
CLPGSVRSSSWVCVDNSVIFRSNSPFPSPRVFMDSNGIRQFCVHVNNNSNLNYFQKLQKVNA  
TNFQALAAEFGGESFTSTFQTQSPPSFYRAGDPILTYFPKWSVISLLRQPAGVGAGGLCA  
ESNPAGFLESKSTTCTRFFKNLASSCTLDSALNAASYNFTVLKVPRSM TDPQNMEFQVP  
VILTSQANAPLLAGNTCQNVVSQVTYE IETNGTFGIQKVS VSLGQTNLTVEPGASLQQHF  
ILRFRAFQQSTAASLTSPRSGNPGYIVGKPLLALTDDISYSMTLLQSQNGSCSVKRHEV  
QFGVNAISGCKLRLKKADC SHLQQEIYQTLHGRPRPEYVAIFGNADPAQKGGWTRILNRH  
CSISAINCTSCCLIPVSLEIQVLWAYVGLLSNPQAHVSGVRFLYQCQSIQDSQQVTEVSL  
TTLVNFVDITQKPQPPRGQPKMDWKWPFDFFPFKVAFSRGVFSQKCSVSPILILCLLLLG  
VLNLETM

**Important features of the protein:****Signal peptide:**

amino acids 1-22

**Transmembrane domains:**

amino acids 484-505, 581-600

**N-glycosylation sites:**amino acids 78-82, 165-169, 179-185, 279-285, 331-337,  
347-351, 410-414, 487-491**N-myristoylation sites:**

amino acids 30-36, 41-47, 124-130, 232-238, 236-242, 409-415

**Prokaryotic membrane lipoprotein lipid attachment site:**

amino acids 420-431

302/392

**FIGURE 289**

CGCGGAGCCCTGCGCTGGGAGGTGCACGGTGTGCACGCTGGACTGGACCCCCATGCAACC  
CCGCGCCCTGCGCCTTAACCAGGACTGCTCCGCGCGCCCCCTGAGCCTCGGGCTCCGGCCCC  
GGACCTGCAGCCTCCCAGGTGGCTGGGAAGAACTCTCCAACAATAAATACATTTGATAAG  
AAAGATGGCTTTAAAAGTGCTACTAGAACAAAGAGAAAAACGTTTTTCACTCTTTTAGTATT  
ACTAGGCTATTTGTGCTAAAGTGACTTGTGAATCAGGAGACTGTAGACAGCAAGAATT  
CAGGGATCGGTCTGGAAGCTGTGTTCCCTGCAACCAGTGTGGGCCAGGCATGGAGTTGTC  
TAAGGAATGTGGCTTCGGCTATGGGGAGGATGCACAGTGTGTGACGTGCCGGCTGCACAG  
GTTCAAGGAGGACTGGGGCTTCAGAAATGCAAGCCCTGTCTGGAAGTGCAGTGGTGAA  
CCGCTTTCAGAAGGCAAATTGTTACGCCACCAAGTGTGCCATCTGCGGGGACTGCTTGCC  
AGGATTTTATAGGAAGACGAACTTGTTCGGCTTTCAAGACATGGAGTGTGTGCCTTGTGG  
AGACCCCTCCTCCTTACGAACCGCACTGTGCCAGCAAGGTCAACCTCGTGAAGATCGC  
GTCCACGGCCTCCAGCCCACGGGACACGGCGCTGGCTGCCGTTATCTGCAGCGCTCTGGC  
CACCGTCTGTGCGCCCTGCTCATCCTCTGTGTGTCATCTATTGTAAGAGACAGTTTATGGA  
GAAGAAACCCAGCTGGTCTCTGCGGTGCGAGGACATTCAGTACAACGGCTCTGAGCTGTC  
GTGTTTTGACAGACCTCAGCTCCACGAATATGCCACAGAGCCTGCTGCCAGTGCCGCCG  
TGAAGTCAAGTGCAGACCTGCGGGCCGGTGCCTTGCTCCCATCCATGTGCTGTGAGGAGGC  
CTGCAGCCCCAACCCGGCGACTCTTGGTTGTGGGGTGCATTCTGCAGCCAGTCTTCAGGC  
AAGAAACGCAGGCCAGCCGGGGAGATGGTGCCGACTTTCTTCGGATCCCTCACGCAGTC  
CATCTGTGGCGAGTTTTCAGATGCCTGGCCTCTGATGCAGAATCCCATGGGTGGTGACAA  
CATCTCTTTTTGTGACTCTTATCCTGAACCTCACTGGAGAAGACATTCATTCTCTCAATCC  
AGAACTTGAAAGCTCAACGTCTTTGGATTCAAATAGCAGTCAAGATTTGGTTGGTGGGGC  
TGTTCCAGTCCAGTCTCATTCTGAAAACCTTTACAGCAGCTACTGATTTATCTAGATATAA  
CAACACACTGGTAGAATCAGCATCAACTCAGGATGCACTAACTATGAGAAGCCAGCTAGA  
TCAGGAGAGTGGCGCTGTCATCCACCCAGCCACTCAGACGTCCCTCCAGGAAGCTTAAAG  
AACCTGCTTCTTTCTGTCAGTAGAAGCGTGTGCTGGAACCCAAAGAGTACTCCTTTGTTAG  
GCTTATGGAGTGAAGCTCTGGACCTTGTCATGGCTTCTGGGGCAAAAATAAATCTGAACC  
AACTGACGGCATTGTAAGCCTTTCAGCCAGTTGCTTCTGAGCCAGACCAGCTGTAAGCT  
GAAACCTCAATGAATAACAAGAAAAGACTCCAGGCCGACTCATGATACTCTGCATCTTTC  
CTACATGAGAAGCTTCTCTGCCACAAAAGTGACTTCAAAGACTGATGGGTTGAGCTGGCA  
GCCTATGAGATTGTGGACATATAACAAGAAACAGAAATGCCCTCATGCTTATTTTCATGG  
TGATTGTGGTTTTACAAGACTGAAGACCCAGAGTATACTTTTCTTTCCAGAAATAATTT  
CATACCGCCTATGAAATATCAGATAAATTACCTTAGCTTTTATGTAGAATGGGTTCAAAA  
GTGAGTGTTCCTATTTGAGAAGGACACTTTTTATCATCTAAACTGATTCGCATAGGTGG  
TTAGAATGGCCCTCATATTGCCTGCCTAAATCTTGGGTTTATTAGATGAAGTTTACTGAA  
TCAGAGGAATCAGACAGAGGAGGATAGCTTTCCAGAATCCACACTTCTGACCTCAGCC  
TCGGTCTCATGAACACCCGCTGATCTCAGGAGAACACCTGGGCTAGGGAATGTGGTTCGAG  
AAAGGGCAGCCCATTGCCCAGAATTAACACATATTGTAGAGACTTGTATGCAAAGGTTGG  
CATATTTATATGAAAATTAGTTGCTATAGAAACATTTGTTGCATCTGTCCCTCTGCCTGA  
GCTTAGAAGGTTATAGAAAAGGGTATTTATAAACATAAATGACCTTTTACTTGCAATTGT  
ATCTTATACTAAAGGCTTTAGAAATTACAACATATCAGGTTCCCCTACTACTGAAGTAGC  
CTTCCGTGAGAACACACCACATGTTAGGACTAGAAGAAAATGCACAATTTGTAGGGGTTT  
GGATGAAGCAGCTGTAAGTCCCTAGTGTAGTTTGACCAGGACATTGTGCTGCTCCTTCC  
AATTGTGTAAGATTAGTTAGCACATCATCTCCTACTTTAGCCATCCGGTGTGGATTAA  
GAGGACGGTGTCTTCTTCTATTAAAGTGCTCCATCCCCTACCATCTACACATTAGCATTG  
TCTCTAGAGCTAAGACAGAAATTAACCCCGTTCAGTCACAAAGCAGGGAATGGTTCAATT  
ACTCTTAATCTTTATGCCCTGGAGAAGACCTACTTGAACAGGGCATATTTTTTAGACTTC  
TGAACATCAGTATGTTTCGAGGGTACTATGATATTTTGGTTTGGAAATTGCCCTGCCCAAGT  
CACTGTCTTTTAACTTTTAACTGAATATTAATAATGTATCTGTCTTTCCT

303/392

**FIGURE 290**

MALKVLLLEQEKTFFTLLVLLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMELSK  
ECGFGYGEDAQCVTCLHRFKEDWGFQKCKPCLDCAVVNRFQKANC SATSDAICGDCLPG  
FYRKTCLVGFQDMECVPCGDP PPPPYEPHCASKVNLVKIASTASSPRDTALAAVICSALAT  
VLLALLILCVIYCKRQFMKKPSWSLRSQDIQYNGSELSCFDRPQLHEYAHRACCQCRD  
SVQTCGPVRLLLPSMCCEEACSPNPATLGCGVHSAASLQARNAGPAGEMVPTFFGSLTQSI  
CGEFSDAWPLMQNPMGGDNISFCDSYPELTGEDIHSLNPELESSTSLDSNSSQDLVGGAV  
PVQSHSENFTAATDLSRYNNTLVESASTQDALTMRSQLDQESGAVIHPATQTSLQEA

**Important features of the protein:****Signal peptide:**

Amino acids 1-25

**Transmembrane domain:**

Amino acids 169-192

**N-glycosylation sites:**

Amino acids 105-109; 214-218; 319-323; 350-354; 368-372; 379-383

**cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

Amino acids 200-204; 238-242

**Tyrosine kinase phosphorylation site:**

Amino acids 207-214

**N-myristoylation sites:**

Amino acids 55-61; 215-221; 270-276

**Prokaryotic membrane lipoprotein lipid attachment site:**

Amino acids 259-270

**TNFR/NGFR family cysteine-rich region proteins:**

Amino acids 89-96

304/392

## FIGURE 291

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCCGGTTCCGCATC  
CAGCCTAGCGTGTCACGATGCGGCTGGGCTCCGGGACTTTTCGCTACCTGTTGCGTAGCG  
ATCGAGGTGCTAGGGATCGCGGTCTTCCTTCGGGGATTCTTCCCGGCTCCCGTTCCGTTCC  
TCTGCCAGAGCGGAACACGGAGCGGAGCCCCAGCGCCCGAACCCTCGGCTGGAGCCAGT  
TCTAACTGGACCACGCTGCCACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGAT  
GCCTTGAGAGATGATTTTGTGTTTGGGTCAAAGGGTGTGAAATTTATGCCCTACACAAC  
TACCTTGTGGAAAAAGGAGCATCTCACAGTTTTGTGGCTGAAGCAAAGCCACCTACAGTT  
ACTATGCCCTCGAATCAAGGCATTGATGACGGGGAGCCTTCCTGGCTTGTGCGACGTCATC  
AGGAACCTCAATTCTCTGCACTGCTGGAAGACAGTGTGATAAGACAAGCAAAGCAGCT  
GGAAAAAGAATAGTCTTTTATGGAGATGAAACCTGGGTAAATTTATCCCAAAGCATT  
GTGGAATATGATGGAACAACCTCATTTTTTCGTGTGAGATTACACAGAGGTGGATAATAAT  
GTCACGAGGCATTTGGATAAAGTATTAAAAAGAGGAGATTGGGACATATAATCCTCCAC  
TACCTGGGGCTGGACCACATTGGCCACATTTTCAGGGCCCAACAGCCCCCTGATTGGGCAG  
AAGCTGAGCGAGATGGACAGCGTGCTGATGAAGATCCACACCTCACTGCAGTCGAAGGAG  
AGAGAGACGCCTTTACCCAATTTGCTGGTTCTTTGTGGTGACCATGGCATGTCTGAAACA  
GGAAGTCACGGGGCCTCCTCCACCGAGGAGGTGAATACACCTCTGATTTTAATCAGTTCT  
GCGTTTGAAAGGAAACCCGGTGATATCCGACATCCAAAGCACGTCCAATAGACGGATGTG  
GCTGCGACACTGGCGATAGCACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTC  
CTATTCCCAGTTGTGGAAGGAAGACCAATGAGAGAGCAGTTGAGATTTTACATTTGAAT  
ACAGTGCAGCTTAGTAACTGTTGCAAGAGAATGTGCCGTGATATGAAAAAGATCCTGGG  
TTTGAGCAGTTTAAAATGTCAGAAAGATTGCATGGGAACTGGATCAGACTGTACTTGGAG  
GAAAAGCATTGAGAAATCCTATTCAACCTGGGCTCCAAGGTTCTCAGGCAGTACCTGGAT  
GCTCTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCCAGTTCTCACCCCTGCTCC  
TGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCCCTGTCTATCTC  
CTGGGTTTTCTCTGCTCTTTTATTTGGTGATCCTGGTTCTTTTCGGCCGTTTACGTCATTG  
TGTGCACCTCAGCTGAAAGTTTCGTGCTACTTCTGTGGCCTCTCGTGGCTGGCGGCAGGCT  
GCCTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGGCAGTGCCC  
TGGACAGGGGGCCTCAGGGAAGGACGTGGAGCAGCCTTATCCCAGGCCTCTGGGTGTCCC  
GACACAGGTGTTACATCTGTGCTGTCAGGTGAGATGCCTCAGTTCTTGAAAGCTAGGT  
TCCTGCGACTGTTACCAAGGTGATTGTAAAGAGCTGGCGGTACAGAGGAACAAGCCCC  
CAGCTGAGGGGGTGTGTGAATCGGACAGCCTCCCAGCAGAGGTGTGGGAGCTGCAGCTGA  
GGGAAGAAGAGACAATCGGCCTGGACACTCAGGAGGGTCAAAGGAGACTTGGTCGCACC  
ACTCATCCTGCCACCCCCAGAATGCATCCTGCCTCATCAGGTCCAGATTTCTTTCCAAGG  
CGGACGTTTTCTGTTGGAATTCTTAGTCTTGGCCTCGGACACCTTCATTCTGTTAGCTGG  
GGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCACACTCAGATCCACAGAGCCCA  
GGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCCCACCCTGCAC  
AGCCCTCATCCCCTCTGGCTTGAGCCGTGAGAGGCCCTGTGCTGAGTGTCTGACCGAGA  
CACTCACAGCTTTGTCATCAGGGCACAGGCTTCCTCGGAGCCAGGATGATCTGTGCCACG  
CTTGACCTCGGGCCCATCTGGGCTCATGCTCTCTCTCTGCTATTGAATTAGTACCTAG  
CTGCACACAGTATGTAGTTACCAAAGAATAAACGGCAATAATTGAGAAAAAAA



305/392

**FIGURE 292**

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTL  
PPPLFSKVIVLIDALRDDVFVFGSKGVKFMPTTYLVEKGASHSFVAEAKPPTVTMPRIK  
ALMTGSLPGFVDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGT  
TSFFVSDYTEVDNNVTRHLDKVLKRGDWDILILHYLGLDGHISGPNSPLIGQKLSEMD  
SVLMKIHTSLQSKERETPLPNLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKP  
GDIRHPKHVQ

**Important features of the protein:****Signal peptide:**

amino acids 1-34

**Transmembrane domain:**

amino acids 58-76

**N-glycosylation sites:**

amino acids 56-60, 194-198

**N-myristoylation sites:**amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,  
275-281, 278-284**Amidation site:**

amino acids 154-158

**Cell attachment sequence:**

amino acids 205-208

306/392

**FIGURE 293**

AGCCAGGCAGCACATCACAGCGGGAGGAGCTGTCCCAGGTGGCCCAGCTCAGCAATGGCA  
ATGGGGGTCCCCAGAGTCATTCTGCTCTGCCTCTTTGGGGCTGCGCTCTGCCTGACAGGG  
TCCAAGCCCTGCAGTGCTACAGCTTTGAGCACACCTACTTTGGCCCCCTTGACCTCAGG  
GCCATGAAGCTGCCCAGCATCTCCTGTCTCATGAGTGCTTTGAGGCTATCCTGTCTCTG  
GACACCGGGTATCGCGCGCCGGTGACCTGGTGCGGAAGGGCTGCTGGACCGGGCCTCCT  
GCGGGCCAGACGCAATCGAACCCGGACGCGCTGCCGCCAGACTACTCGGTGGTGCGCGGC  
TGCACAACTGACAAATGCAACGCCACCTCATGACTCATGACGCCCTCCCCAACCTGAGC  
CAAGCACCCGACCCGCCGACGCTCAGCGGCGCCGAGTGCTACGCCTGTATCGGGGTCCAC  
CAGGATGACTGCGCTATCGGCAGGTCCCGACGAGTCCAGTGTCACCAGGACCAGACCGCC  
TGCTTCCAGGGCAGTGGCAGAATGACAGTTGGCAATTTCTCAGTCCCTGTGTACATCAGA  
ACCTGCCACCGGCCCTCCTGCACCACCGAGGGCACCACCAGCCCCCTGGACAGCCATCGAC  
CTCCAGGGCTCCTGCTGTGAGGGGTACCTCTGCAACAGGAAATCCATGACCCAGCCCTTC  
ACCAGTGCTTCAGCCACCACCCCTCCCCGAGCACTACAGGTCCCTGGCCCTGCTCCTCCCA  
GTCCTCCTGCTGGTGGGGCTCTCAGCATAGACCGCCCCTCCAGGATGCTGGGGACAGGGC  
TCACACACCTCATTCTTGCTGCTTCAGCCCCTATCACATAGCTCACTGGAAAATGATGTT  
AAAGTAAGAATTGCAAAA

307/392

**FIGURE 294**

MAMGVPRVILLCLFGAALCLTGSQALQCYSFEHTYFGPFDLRAMKLPSISCPHECFEAIL  
SLDTGYRAPVTILVRKGCWTGPPAGQTQSNPDALPPDYSVVRGCTTDKCNAHLMTHDALPN  
LSQAPDPPTLSGAECYACIGVHQDDCAIGRSRRVQCHQDQTACFQSGRMTVGNFSPVY  
IRTCHRPSCCTTEGTTSPWTAIDLQSCCEGYLCNRKSMTQPFTSASATTPPRALQVLALL  
LPVLLLVGLSA

**Important features of the protein:**

**Signal peptide:**

amino acids 1-19

**Transmembrane domain:**

amino acids 233-251

**N-glycosylation sites:**

amino acids 120-124, 174-178

**N-myristoylation sites:**

amino acids 15-21, 84-90

308/392

**FIGURE 295**

AATCGGCTGATTCTGCATCTGGAACTGCCTTCATCTTGAAAGAAAAGCTCCAGGTCCCT  
TCTCCAGCCACCCAGCCCCAAGATGGTGATGCTGCTGCTGCTGCTTTCCGCACTGGCTGG  
CCTCTTCGGTGCGGCAGAGGGACAAGCATTTTCATCTTGGGAAGTGCCCCAATCCTCCGGT  
GCAGGAGAATTTTGACGTGAATAAGTATCTCGGAAGATGGTACGAAATTGAGAAGATCCC  
AACAAACCTTTGAGAATGGACGCTGCATCCAGGCCAACTACTCACTAATGGAAAACGGAAA  
GATCAAAGTGTTAAACCAGGAGTTGAGAGCTGATGGAAGTGAATCAAATCGAAGGTGA  
AGCCACCCAGTTAACCTCACAGAGCCTGCCAAGCTGGAAGTTAAGTTTTCTGGTTTTAT  
GCCATCGGCACCGTACTGGATCCTGGCCACCGACTATGAGAAGTATGCCCTCGTGTATTC  
CTGTACCTGCATCATCCAACTTTTTCACGTGGATTTTGCTTGGATCTTGGCAAGAAACCC  
TAATCTCCCTCCAGAAACAGTGGACTCTCTAAAAAATATCCTGACTTCTAATAACATTGA  
TGTCAGAAAATGACGGTCACAGACCAGGTGAACTGCCCCAAGCTCTCGTAAACAGGTTTC  
TACAGGGAGGCTGCACCCACTCCATGTTACTTCTGCTTCGCTTTCCCCTACCCACCCCC  
CCCCATAAGACAAACCAATCAACCACGACAAAGGAAGTTGACCTGAACATGTAACCAT  
GCCCTACCCTGTTACCTTGCTAGCTGCAAAATAAACTTGTTGCTGACCTGCTGTGCTCGC  
AAAAAA

309/392

## FIGURE 296

MVMLLLLLLSALAGLFGAAEGQAFHLGKCPNPPVQENFDVNKYLGRWYEIEKIPITTFENG  
RCIQANYSLMENGKIKVINQELRADGTVNQIEGEATPVNLTEPAKLEVKFSWFMPSPAY  
WILATDYENYALVYSCTCIIQLFHVDFAWILARNPNLPPETVDSLKNILTSNNIDVKKM  
TVTDQVNCPKLS

Signal sequence:

1-16

N-glycosylation site:

65-68

98-101

cAMP- and cGMP-dependent protein kinase phosphorylation  
site:

175-178

N-myristoylation site:

13-18

16-21

Lipocalin proteins:

36-47

120-130

Lipocalin / cytosolic fatty-acid binding proteins:

41-185

310/392

**FIGURE 297**

GGGTGATTGAACTAAACCTTCGCCGCACCGAGTTTGACGTACGGCCGTCACCCGCACCGC  
TGCCTGCTTGCGGTTGGAGAAATCAAGGCCCTACCGGGCCTCCGTAGTCACCTCTCTATA  
GTGGGCGTGGCCGAGGCCGGGGTGACCCTGCCGGAGCCTCCGCTGCCAGCGACATGTTCA  
AGGTAATTCAGAGGTCCGTGGGGCCAGCCAGCCTGAGCTTGCTCACCTTCAAAGTCTATG  
CAGCACCAAAAAAGGACTCACCTCCCAAAAATTCCGTGAAGGTTGATGAGCTTTCACTCT  
ACTCAGTTCTTGAGGGTCAATCGAAGTATGTGGAGGAGGCAAGGAGCCAGCTTGAAGAAA  
GCATCTCACAGCTCCGACACTATTGCGAGCCATACACAACCTGGTGTGAGGAAACGTACT  
CCCAAACCTAAGCCCAAGATGCAAAGTTTGGTTCAATGGGGGTTAGACAGCTATGACTATC  
TCCAAAATGCACCTCCTGGATTTTTTCCGAGACTTGGTGTATTGGTTTTGCTGGCCTTA  
TTGGACTCCTTTTGGCTAGAGGTTCAAAAATAAAGAAGCTAGTGTATCCGCCTGGTTTTCA  
TGGGATTAGCTGCCTCCCTCTATTATCCACAACAAGCCATCGTGTTTGCCAGGTCAGTG  
GGGAGAGATTATATGACTGGGGTTTACGAGGATATATAGTCATAGAAGATTTGTGGAAGG  
AGAACTTTCAAAGCCAGGAAATGTGAAGAATTCACCTGGAACCTAAGTAGAAAACCTCCAT  
GCTCTGCCATCTTAATCAGTTATAAGGTAAACATTGGAAACTCCATAGAATAAATCAGTAT  
TTCTACAGAAAAATGGCATAGAAGTCAGTATTGAATGTATTAAATTGGCTTTCTTCTTCA  
GGAAAACTAGACCAGACCTCTGTTATCTTCTGTGAAATCATCCTACAAGCAAACCTAACC  
TGGAATCCCTTCACCTAGAGATAATGTACAAGCCTTAGAACTCCTCATTCTCATGTTGCT  
ATTTATGTACCTAATTAAAACCAAGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAA

311/392

## FIGURE 298

MFKVIQRSVGPASLSLLTFKVYAAPKKDSPPKNSVKVDELSTLYSVPEGQSKYVEEARSQLE  
EESISQLRHYCEPYTTWCQETYSQTKPKMQSLVQWGLDSYDYLQNAPPGFFPRLGVIGFA  
GLIGLLLARGSKIKKLVYPPGFMGLAASLYYPQQAIVFAQVSGERLYDWGLRGYIVIEDL  
WKENFQKPGNVKNSPGTK

**Important features:**

**Signal peptide:**

Amino acids 1-23

**Transmembrane domain:**

Amino acids 111-130

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 26-30

**Tyrosine kinase phosphorylation site:**

Amino acids 36-44

**N-myristoylation sites:**

Amino acids 124-130;144-150;189-195

312/392

**FIGURE 299**

CCGCTGAGATGTACGAACTTCCGGTTCTCCGGGCAGCTGCCACTGCTGTAGCTTCTGCCA  
CCTGCCACGACCGGGCCTCTCCCTGGCGTTTGGTCACCTCTGCTTCATTCTCCACCGCGC  
CTATGGTCCCTCTTGGAGCCAGCGTGGCGGGCCTGGCGGCTCCCGGGTGGTGAGAGAGCG  
GTCCGGGAACGATGAAGGCCTCGCAGTGCTGCTGCTGTCTCAGCCACCTCTTGGCTTCCG  
TCCTCCTCCTGCTGTTGCTGCCTGAACTAAGCGGGCCCCTGGCAGTCCTGCTGCAGGCAG  
CCGAGGCCGCGCCAGGTCTTGGGCCTCCTGACCCTAGACCACGGACATTACCGCCGCTGC  
CACCGGGCCCTACCCCTGCCCAGCAGCCGGGCGGTGGTCTGGCTGAAGCTGCGGGGCCG  
GGGGCTCCGAGGGAGGCAATGGCAGCAACCCTGTGGCCGGGCTTGAGACGGACGATCACG  
GAGGGAAGGCCGGGAAGGCTCGGTGGGTGGCGGCCCTTGCTGTGAGCCCCAACCTGGCG  
ACAAGCCCATGACCCAGCGGGCCCTGACCGTGTGATGGTGGTGAGCGGCGCGGTGCTGG  
TGTACTTTCGTGGTCAGGACGGTCAGGATGAGAAGAAGAAACCGAAAGACTAGGAGATATG  
GAGTTTTGGACACTAACATAGAAAATATGGAATTGACACCTTTAGAACAGGATGATGAGG  
ATGATGACAACACGTTGTTTGATGCCAATCATCCTCGAAGATAAGAATGTGCCTTTTGAT  
GAAAGAAC'TTTATCTTTCTACAATGAAGAGTGGAATTTCTATGTTTAAGGAATAAGAAGC  
CACTATATCAATGTTGGGGGGGTATTTAAGTTACATATATTTTAACAACCTTTAATTTGC  
TGTTGCAATAAATACCGTATCCTTTTATTATATCTTTATATGTATAGAAGTACTCTATTA  
ATGGGCTCAGAGATGTTGGGGATAAAGTATACTGTAATAATTTATCTGTTTGAAAAATTAC  
TATAAAACGGTGTTTTCTGGTGGTTTTTGTTCCTGCTTACCATATGATTGTAAATTGT  
TTTATGTATTAATCAGTTAATGCTAATTATTTTGTCTGATGTCATATGTTAAAGAGCTAT  
AAATTCCAACAACCAACTGGTGTGTAATAATAATTTAAAATTTCCTTTACTGAAAGGTAT  
TTCCCATTTTTGTGGGGAAGAAGCCAAATTTATTACTTTGTGTTGGGGTTTTTAAAT  
ATTAAGAAATGTCTAAGTTATTGTTTGCAAAACAATAAATATGATTTTAAATTCTCTAA  
AAAAAAA



313/392

## FIGURE 300

MKASQCCCCLSHLLASVLLLLLLPELSGPLAVLLQAEEAAPGLGPPDPRPRTLPPPLPPGP  
TPAQQPGRGLAEAAAGPRGSEGGNGSNPVAGLETDDHGGKAGEGSGVGGGLAVSPNPGDKPM  
TQRALTVLMVVSGAVLVYFVVRTVRMRRNRKTRRYGVLDTNIE NMELTPLEQDDEDDDN  
TLFDANHPRR

**Signal peptide:**  
amino acids 1-28

**Transmembrane domain:**  
amino acids 124-140

**N-glycosylation site:**  
amino acids 83-87

**N-myristoylation sites:**  
amino acids 69-75, 78-84, 81-87, 97-103, 103-109, 106-112,  
157-160

314/392

**FIGURE 301**

CTCGGCTGGATTTAAGGTTGCCGCTAGCCGCTGGGAATTTAAGGGACCCACACTACCTT  
CCCGAAGTTGAAGGCAAGCGGTGATTGTTTTGTAGACGGCGCTTTGTCATGGGACCTGTGC  
GGTTGGGAATATTGCTTTTCCTTTTTTTGGCCGTGCACGAGGCTTGGGCTGGGATGTTGA  
AGGAGGAGGACGATGACACAGAACGCTTGCCAGCAAATGCGAAGTGTGTAAGCTGCTGA  
GCACAGAGCTACAGGCGGAAGTGAAGTGCACCGGTGATCTCGAGAGGTGCTGGAGCTGG  
GGCAGGTGCTGGATACAGGCAAGAGGAAGAGACACGTGCCTTACAGCGTTTCAGAGACAA  
GGCTGGAAGAGGCCCTTAGAGAATTTATGTGAGCGGATCCTGGACTATAGTGTTACGCTG  
AGCGCAAGGGCTCACTGAGATATGCCAAGGGTCAGAGTCAGACCATGGCAACACTGAAAG  
GCCTAGTGAGCAAGGGGGTGAAGGTGGATCTGGGGATCCCTCTGGAGCTTTGGGATGAGC  
CCAGCGTGGAGGTACATACCTCAAGAAGCAGTGTGAGACCATGTTGGAGGAGTTTGAAG  
ACATTGTGGGAGACTGGTACTTCCACCATCAGGAGCAGCCCCTACAAAATTTTCTCTGTG  
AAGGTCATGTGCTCCAGCTGCTGAAACTGCATGTCTACAGGAACTTGGACTGGAAAGG  
AGATCACAGATGGGGAAGAGAAAACAGAAGGGGAGGAAGAGCAGGAGGAGGAGGAGGAAG  
AGGAGGAAGAGGAAGGGGGAGACAAGATGACCAAGACAGGAAGCCACCCCAAACCTTGACC  
GAGAAGATCTTTGACCCCTTGCTTTGAGCCCCCAGGAGGGGAAGGGATCATGGAGAGCCC  
TCTAAAGCCTGCACCTCTCCCTGCTCCACAGCTTTCAGGGTGTGTTTATGAGTGAATCCAC  
CCAAGCTTGTAGCTGTTCTCTCCCATCTAACCTCAGGCAAGATCCTGGTGAAACAGCATG  
ACATGGCTTCTGGGGTGGAGGGTGGGGGTGGAGGTCTGCTCCTAGAGATGAACTCTATC  
CAGCCCCCTTAATTGGCAGGTGTATGTGCTGACAGTACTGAAAGCTTTCCTCTTTAACTGA  
TCCCCCCCCACCCAAAAGTCAGCAGTGGCACTGGAGCTGTGGGCTTTGGGGAAGTCACT  
TAGCTCCTTAAGGTCTGTTTTTAGACCCTTCCAAGGAAGAGGCCAGAACGGACATTCTCT  
GCGATCTATATACATTGCCTGTATCCAGGAGGCTACACACCAGCAAACCGTGAAGGAGAA  
TGGGACACTGGGTGATGGCCTGGAGTTGCTGATAATTTAGGTGGGATAGATACTTGGTCT  
ACTTAAGCTCAATGTAACCCAGAGCCCACCATATAGTTTTATAGGTGCTCAACTTTCTAT  
ATCGCTATTAACTTTTTTCTTTTTTTCTA

315/392

**FIGURE 302**

MGPVRLGILLFLFLAVHEAWAGMLKEEDDDTERLPSKCEVCKLLSTELQAELSRTGRSRE  
VLELGQVLDTGKRKRHPYPSVSETRLEEALENLCERILDYSVHAERKGSRLRYAKGQSQTM  
ATLKGLVQKGVKVDLGIPLELWDEPSVEVTYLKKQCETMLEEFEDIVGDWYFHHQEQLQ  
NFLCEGHVLPAAETACLQETWTGKEITDGEEKTEGEEEQEEEEEEEEEGGDKMTKTGSH  
PKLDREDL

**Important features of the protein:**

**Signal peptide:**

amino acids 1-21

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

amino acids 106-110

**N-myristoylation site:**

amino acids 115-121

**Amidation site:**

amino acids 70-74

316/392

**FIGURE 303**

CTCCTGCACTAGGCTCTCAGCCAGGGATGATGCGCTGCTGCCGCCGCCGCTGCTGCTGCC  
GGCAACCACCCCATGCCCTGAGGCCGTTGCTGTTGCTGCCCCCTCGTCCTTTTACCTCCCC  
TGGCAGCAGCTGCAGCGGGCCCAAACCGATGTGACACCATATACCAGGGCTTCGCCGAGT  
GTCTCATCCGCTTGGGGGACAGCATGGGCCGCGGAGGCGAGCTGGAGACCATCTGCAGGT  
CTTGGAATGACTTCCATGCCTGTGCCTCTCAGGTCCTGTCAGGCTGTCCGGAGGAGGCAG  
CTGCAGTGTGGGAATCACTACAGCAAGAAGCTCGCCAGGCCCCCCCGTCCGAATAACTTGC  
ACACTCTGTGCGGTGCCCCGGTGCAATGTTGCGGAGCGCGGCACAGGCTCCGAAACCAACC  
AGGAGACGCTGCGGGCTACAGCGCCTGCACTCCCCATGGCCCCCTGCGCCCCCACTGCTGG  
CGGCTGCTCTGGCTCTGGCCTACCTCCTGAGGCCTCTGGCCTAGCTTGTTGGGTTGGGTA  
GCAGCGCCCGTACCTCCAGCCCTGCTCTGGCGGTGGTTGTCCAGGCTCTGCAGAGCGCAG  
CAGGGCTTTTCATTAAAGGTATTTATATTTGTA

317/392

## FIGURE 304

MMRCCRRRCCCRQPPHALRPLLLLPLVLLPPLAAAAAGPNRCDTIYQGFAECLIRLGDSM  
GRGGELETICRSWNDFHACASQVLSGCPEEAAAVWESLQQEARQAPRPNNLHTLCGAPVH  
VRERGTGSETNQETLRATAPALPMAPAPPLAAAALALAYLLRPLA

**Signal peptide:**

Amino acids 1-35

**Transmembrane domain:**

Amino acids 141-157

**N-myristoylation site:**

Amino acids 127-133

**Prokaryotic membrane lipoprotein lipid attachment site:**

Amino acids 77-88

318/392

**FIGURE 305**

AAGTACTTGTGTCCGGGTGGTGGACTGGATTAGCTGCGGAGCCCTGGAAGCTGCCTGTCC  
TTCTCCCTGTGCTTAACCAGAGGTGCCCATGGGTTGGACAATGAGGCTGGTCACAGCAGC  
ACTGTTACTGGGTCTCATGATGGTGGTCACTGGAGACGAGGATGAGAACAGCCCGTGTGC  
CCATGAGGCCCTCTTGGACGAGGACACCCTCTTTTGCCAGGGCCTTGAAGTTTTCTACCC  
AGAGTTGGGGAACATTGGCTGCAAGGTTGTTCCCTGATTGTAACAACCTACAGACAGAAGAT  
CACCTCCTGGATGGAGCCGATAGTCAAGTTCCCGGGGGCCGTGGACGGCGCAACCTATAT  
CCTGGTGATGGTGGATCCAGATGCCCCCTAGCAGAGCAGAACCAGACAGAGATTCTGGAG  
ACATTGGCTGGTAACAGATATCAAGGGCGCCGACCTGAAGAAAGGAAGATTCAAGGCCA  
GGAGTTATCAGCCTACCAGGCTCCCTCCCCACCGGCACACAGTGGCTTCCATCGCTACCA  
GTTCTTTGTCTATCTTCAGGAAGGAAAAGTCATCTCTCTCTCCTTCCCAAGGAAAACAAAAC  
TCGAGGCTCTTGGAAAATGGACAGATTTCTGAACCGCTTCCACCTGGGCGAACCTGAAGC  
AAGCACCAGTTTCATGACCCAGAACTACCAGGACTACCAACCCTCCAGGCTCCAGAGG  
AAGGGCCAGCGAGCCCAAGCACAAAACCAGGCAGAGATAGCTGCCTGCTAGATAGCCGGC  
TTTGCCATCCGGGCATGTGGCCACACTGCTCACCACCGACGATGTGGGTATGGAACCCCC  
TCTGGATACAGAACCCTTCTTTTCCAAATTAATAAAAAAAAAAATCATCAA

319/392

**FIGURE 306**

MGWTMRLVTAALLLGLMMVVTGDEDENSPCAHEALLDEDTLFCQGLEVFYPELGNIGCKV  
VPDCNNYRQKITSWMEPIVKFPGAVDGATYILVMVDPDAPSRAEPRQRFWRHWLVTDIKG  
ADLKKGKIQQEELSAYQAPSPPAHSGFHRYQFFVYLQEGKVISLLPKENKTRGSWKMDRF  
LNRFHLGEPEASTQFMTQNYQDSPTLQAPRGRASEPKHKTRQR

**Important features of the protein:**

**Signal peptide:**

amino acids 1-22

**N-glycosylation site:**

amino acids 169-173

**Tyrosine kinase phosphorylation site:**

amino acids 59-68

**N-myristoylation sites:**

amino acids 54-60, 83-89, 130-136

**Phosphatidylethanolamine signature:**

amino acids 113-157

320/392

**FIGURE 307**

AAGGAGCAGCCCGCAAGCACCAAGTGAGAGGCATGAAGTTACAGTGTGTTTCCCTTTGGC  
TCCTGGGTACAATACTGATATTGTGCTCAGTAGACAACCACGGTCTCAGGAGATGTCTGA  
TTTCCACAGACATGCACCATATAGAAGAGAGTTTCCAAGAAATCAAAGAGCCATCCAAG  
CTAAGGACACCTTCCCAAATGTCACATCCTGTCCACATTGGAGACTCTGCAGATCATT  
AGCCCTTAGATGTGTGCTGCGTGACCAAGAACCTCCTGGCGTTCTACGTGGACAGGGTGT  
TCAAGGATCATCAGGAGCCAAACCCCAAATCTTGAGAAAAATCAGCAGCATTGCCAACT  
CTTTCCTCTACATGCAGAAAACCTGCGGCAATGTGAGAACAGAGGCAGTGTCACTGCA  
GGCAGGAAGCCACCAATGCCACCAGAGTCATCCATGACAACCTATGATCAGCTGGAGGTCC  
ACGCTGCTGCCATTAAATCCCTGGGAGAGCTCGACGTCTTTCTAGCCTGGATTAATAAGA  
ATCATGAAGTAATGTTCTCAGCTTGATGACAAGGAACCTGTATAGTGATCCAGGGATGAA  
CACCCCCTGTGCGGTTTACTGTGGGAGACAGCCACCTTGAAGGGGAAGGAGATGGGGAA  
GGCCCCCTTGACAGCTGAAAGTCCCACTGGCTGGCCTCAGGCTGTCTTATTCCGCTTGAAAA  
TAGGCAAAAAGTCTACTGTGGTATTTGTAATAAACTCTATCTGCTGAAAGGGCCTGCAGG  
CCATCCTGGGAGTAAAGGGCTGCCTTCCCATCTAATTTATTGTAAAGTCATATAGTCCAT  
GTCTGTGATGTGAGCCAAGTGATATCCTGTAGTACACATTGTACTGAGTGGTTTTTCTGA  
ATAAATTCCATATTTTACCTATGA



321/392

**FIGURE 308**

MKLQCVSLWLLGTILILCSVDNHGLRRCLISTDMHHIEESFQEIKRAIQAKDTFPNVTIL  
STLETLQIIKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSFLYMQKTLRQ  
CQEQRQCHCRQEATNATRVIHNDYDQLEVHAAAIKSLGELDVFLAWINKNHEVMFSA

**Signal sequence:**

amino acids 1-18

**N-glycosylation sites:**

amino acids 56-60, 135-139

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

amino acids 102-106

**N-myristoylation site:**

amino acids 24-30

**Actinin-type actin-binding domain signature 1:**

amino acids 159-169

322/392

**FIGURE 309**

GTCGACCCACGCGTCCGAAGCTGCTGGAGCCACGATTTCAGTCCCCTGGACTGTAGATAAA  
GACCCTTCTTGCCAGGTGCTGAGACAACCACACTATGAGAGGCACTCCAGGAGACGCTG  
ATGGTGGAGGAAGGGCCGTCTATCAATCAATCACTGTTGCTGTTATCACATGCAAGTATC  
CAGAGGCTCTTGAGCAAGGCAGAGGGGATCCCATTTATTTGGGAATCCAGAATCCAGAAA  
TGTGTTTGTATTGTGAGAAGGTTGGAGAACAGCCACATTGCAGCTAAAAGAGCAGAAGA  
TCATGGATCTGTATGGCCAACCCGAGCCCGTGAAACCCCTCCTTTTCTACCGTGCCAAGA  
CTGGTAGGACCTCCACCCTTGAGTCTGTGGCCTTCCCGGACTGGTTCATTGCCTCCTCCA  
AGAGAGACCAGCCCATCATTCTGACTTCAGAACTTGGGAAGTCATACAACACTGCCTTTG  
AATTAAATATAAATGACTGA<sup>1</sup>ACTCAGCCTAGAGGTGGCAGCTTGGTCTTTGTCTTAAAGT  
TTCTGGTTCCCAATGTGTTTTCGTCTACATTTTCTTAGTGTCAATTTTCACGCTGGTGCTG  
AGACAGGAGCAAGGCTGCTGTTATCATCTCATTTTATAATGAAGAAGAAGCAATTACTTC  
ATAGCAACTGAAGAACAGGATGTGGCCTCAGAAGCAGGAGAGCTGGGTGGTATAAGGCTG  
TCCTCTCAAGCTGGTGCTGTGTAGGCCACAAGGCATCTGCATGAGTGACTTTAAGACTCA  
AAGACCAAACACTGAGCTTTCTTCTAGGGGTGGGTATGAAGATGCTTCAGAGCTCATGCG  
CGTTACCCACGATGGCATGACTAGCACAGAGCTGATCTCTGTTTCTGTTTTGCTTTATTC  
CCTCTTGGGATGATATCATCCAGTCTTTATATGTTGCCAATATACCTCATTGTGTGTAAT  
AGAACCTTCTTAGCATTAAAGACCTTGTAACAAAAATAATTCCTGGGGTGGGTATGAAGA  
TGCTTCAGAGCTCATGCGCGTTACCCACGATGGCATGACTAGCACAGAGCTGATCTCTGT  
TTCTGTTTTGCTTTATCCCTCTTGGGATGATATCATCCAGTCTTTATATGTTGCCAATA  
TACCTCATTGTGTGTAATAGAACCTTCTTAGCATTAAAGACCTTGTAACAAAAATAATTC  
TTGTGTAAAGTTAAATCATTTTTGTCTAATTGTAATGTGTAATCTTAAAGTTAAATAAA  
CTTTGTGTATTTATATAATAATAAGCTAA<sup>2</sup>ACTGATATAAAATAAAGAAAGAGTAAACTG

323/392

## FIGURE 310

MRGTPGDADGGGRAVYQSITVAVITCKYPEALEQGRGDPIYLGIONPEMCLYCEKVGEQP  
TLQLKEQKIMDLYGQPEPVKPFIFYRAKTGRTSTLESVAFPDWFIASSKRDQPIILTSEL  
GKSYNTAFELNIND

Signal sequence:  
amino acids 1-17

N-myristoylation site:  
amino acids 10-16

Cell attachment sequence:  
amino acids 36-39

324/392

**FIGURE 311**

GCGAGGCTGCACCAGCGCCTGGCACCATGAGGACGCCTGGGCCTCTGCCCCGTGCTGCTGC  
TGCTCCTGGCGGGAGCCCCGCGCGCGGCCCACTCCCCGACCTGCTACTCCCGCATGC  
GGGCCCTGAGCCAGGAGATCACCCGCGACTTCAACCTCCTGCAGGTCTCGGAGCCCTCGG  
AGCCATGTGTGAGATACCTGCCCAGGCTGTACCTGGACATACACAATTACTGTGTGCTGG  
ACAAGCTGCGGGACTTTGTGGCCTCGCCCCGTGTTGGAAAGTGGCCCAGGTAGATTTCCT  
TGAAGGACAAAGCACGGAAGCTGTACACCATCATGAACTCGTTCTGCAGGAGAGATTGG  
TATTCCTGTTGGATGACTGCAATGCCTTGGAATACCCAATCCCAGTGACTACGGTCCTGC  
CAGATCGTCAGCGCTTAAGGGAAGTGAAGACCAGAGAAAGAACCCAAGAGAACTAAAGTTAT  
GTCAGCTACCCAGACTTAATGGGCCAGAGCCATGACCCTCACAGGTCTTGTTAGTTGT  
ATCTGAAACTGTTATGTATCTCTACCTTCTGGAAAACAGGGCTGGTATTCCTACCCAG  
GAACCTCCTTTGAGCATAGAGTTAGCAACCATGCTTCTCATTCCCTTGACTCATGTCTTG  
CCAGGATGGTTAGATACACAGCATGTTGATTTGGTCACTAAAAAGAAGAAAGGACTAAC  
AAGCTTCACTTTTATGAACAACTATTTTGAGAACATGCACAATAGTATGTTTTTATTACT  
GGTTTAATGGAGTAATGGTACTTTTTATTCTTTCTTGATAGAAACCTGCTTACATTTAACC  
AAGCTTCTATTATGCCTTTTTCTAACACAGACTTTCTTCACTGTCTTTCATTTAAAAAGA  
AATTAATGCTCTTAAGATATATATTTTACGTAGTGCTGACAGGACCCACTCTTTCATTGA  
AAGGTGATGAAAATCAAATAAAGAATCTCTTCACATGGA

325/392

## FIGURE 312

MRTPGPLPVLLLLLAGAPAARPTPPTCYSRMRALSQEITRDFNLLQVSEPSEPCVRYLPR  
LYLDIHNYCVLDKLRDFVASPPCWKVAQVDSLKDKARKLYTIMNSFCRRDLVFLDDCNA  
LEYPIPVTTVLPDRQR

Important features of the protein:

Signal peptide:

amino acids 1-19

Tyrosine kinase phosphorylation site:

amino acids 60-69

N-myristoylation site:

amino acids 16-22

326/392

**FIGURE 313**

GAGCGACGCTGTCTCTAGTCGCTGATCCCAATGCACCGGCTCATCTTTGTCTACACTCT  
AATCTGCGCAAACCTTTTGCAGCTGTCTGGGACACTTCTGCAACCCCGCAGAGCGCATCCAT  
CAAAGCTTTGCGCAACGCCAACCTCAGGCGAGATGACTTGTACCGAAGAGATGAGACCAT  
CCAGGTGAAAGGAAACGGCTACGTGCAGAGTCCTAGATTCCCGAACAGCTACCCCAAGAA  
CCTGCTCCTGACATGGCGGCTTCACTCTCAGGAGAATACACGGATACAGCTAGTGTGTTGA  
CAATCAGTTTGGATTAGAGGAAGCAGAAAATGATATCTGTAGGTATGATTTTGTGGAAGT  
TGAAGATATATCCGAAACCAGTACCATTATTAGAGGACGATGGTGTGGACACAAGGAAGT  
TCCTCCAAGGATAAAATCAAGAACGAACCAAATTAATACATTCAAGTCCGATGACTA  
CTTTGTGGCTAAACCTGGATTCAAGATTTATTATTCTTTGCTGGAAGATTTCCAACCCGC  
AGCAGCTTCAGAGACCAACTGGGAATCTGTCACAAGCTCTATTTCAAGGGGTATCCTATAA  
CTCTCCATCAGTAACGGATCCCACTCTGATTGCGGATGCTCTGGACAAAAAATTCAGAG  
ATTTGATACAGTGAAGATCTGCTCAAGTACTTCAATCCAGAGTCATGGCAAGAAGATCT  
TGAGAATATGTATCTGGACACCCCTCGGTATCGAGGCAGGTCATACCATGACCGGAAGTC  
AAAAGTTGACCTGGATAGGCTCAATGATGATGCCAAGCGTTACAGTTGCACTCCCAGGAA  
TTACTCGGTCAATATAAGAGAAGAGCTGAAGTTGGCCAATGTGGTCTTCTTTCCACGTTG  
CCTCCTCGTGCAGCGCTGTGGAGGAAATTGTGGCTGTGGAAGTGTCAACTGGAGGTCCTG  
CACATGCAATTCAGGGAAAACCGTGAAAAAGTATCATGAGGTATTACAGTTTGAGCCTGG  
CCACATCAAGAGGAGGGGTAGAGCTAAGACCATGGCTCTAGTTGACATCCAGTTGGATCA  
CCATGAACGATGCGATTGTATCTGCAGCTCAAGACCACCTCGATAAGAGAATGTGCACAT  
CCTTACATTAAGCCTGAGAGAA

327/392

**FIGURE 314**

MHRLIFVYTLCANFCSCRDTSATPQSASIKALRNANLRRDDLRYRDETIQVKNGYVQS  
PRFPNSYPRNLLLTWRLHSQENTRIQLVFDNQFGLLEEAENDICRYDFVEVEDISETSTII  
RGRWCGHKEVPPRIKSRTNQIKITFKSDDYFVAKPGFKIYYSLLEDFQPAAASETNWESV  
TSSISGVSYNPSVTDPTLIADALDKKIAEFDTVEDLLKYFNPESWQEDLENMYLDTPRY  
RGRSYHDRKSKVDLDRLNDDAKRYSCTPRNYSVNIREELKLANVVFFPRCLLVQRCGGNC  
GCGTVNWRSTCNSGKTVKKYHEVLQFEPGHIKRRGRAKTMALVDIQLDHHERCDCICSS  
RPPR

**Signal peptide:**  
amino acids 1-18

**N-glycosylation site:**  
amino acids 270-274

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**  
amino acids 262-266

**Tyrosine kinase phosphorylation site:**  
amino acids 256-265

**N-myristoylation sites:**  
amino acids 94-100, 186-192, 297-303, 298-304

**TonB-dependent receptor proteins signature 1:**  
amino acids 1-56

328/392

## FIGURE 315

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCTGTGGAG  
CTCAAGATGGTCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTG  
CTTTATCTGCATAATAACCAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCATTAAA  
GGTGAAGAGATCAGCGTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCGTCATC  
CTGGGTGTCCAGGGTGAAGCCAGTGCCTGTCTGTGGGGTGGGGCAGGAGCCGACTCTA  
ACACTAGAGCCAGTGAACATCATGGAGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTC  
ACCTTCTACCGGCGGGACATGGGGCTCACCTCCAGCTTCGAGTCGGCTGCCTACCCGGGC  
TGGTTCCTGTGCACGGTGCCTGAAGCCGATCAGCCTGTGAGCTCACCCAGCTTCCCGAG  
AATGGTGGCTGGAATGCCCCCATCACAGACTTCTACTTCCAGCAGTGTGACTAGGGCAAC  
GTGCCCCCAGAACTCCCTGGGCAGAGCCAGCTCGGGTGAGGGGTGAGTGGAGGAGACCC  
ATGGCGGACAATCACTCTCTCTGCTCTCAGGACCCCCACGTCTGACTTAGTGGGCACCTG  
ACCATTCTGCTCTTCTGGTTCCAGTTTGGATAAAATCTGAGATTGGAGCTCAGTCCACG  
GTCCCTCCCCACTGGATGGTGCTACTGCTGTGGAACCTTGTA AAAACCATGTGGGGTAAA  
CTGGGAATAACATGAAAAGATTCTGTGGGGGTGGGGTGGGGGAGTGGTGGGAATCATTC  
CTGCTTAATGGTAAGTGAACAAGTGTACCCCTGAGCCCCGAGGCCAACCCATCCCCAGTT  
GAGCCTTATAGGGTCACTAGCTCTCCACATGAAGTCCTGTCACTACCACTGTGCAGGAG  
AGGGAGGTGGTCATAGAGTCAGGGATCTATGGCCCTTGGCCAGCCCCACCCCTTCCCT  
TTAATCCTGCCACTGTCTATGCTACCTTTCCCTATCTCTTCCCTCATCATCTTGTGTGG  
GCATGAGGAGGTGGTGATGTCAGAAGAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTA  
GGGTATGCTGATCCTCTTTTAAAAACCCAAGATACAATCAAAATCCCAGATGCTGGTCTC  
TATTTCCCATGAAAAAGTGCTCATGACATATTGAGAAGACCTACTTACAAAGTGGCATATA  
TTGCAATTTATTTTAAATTA AAAAGATACCTATTTATATATTTCTTTATAGAAAAAGTCTG  
GAAGAGTTTACTTCAATTGTAGCAATGTGAGGGTGGTGGCAGTATAGGTGATTTTCTTT  
TAATTCTGTAAATTTATCTGTATTTCTTAATTTTCTACAATGAAGATGAATTCCTGTGA  
TAAAAATAAGAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGATTGTC  
CTCAGCCTCCACTTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTTGG  
TTGTAGTAGTGATCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAG  
TTTGTGTGGCTGGAATCTCTGGGTAAAGGAATTAAGAACA AAAATCATCTGGTAATCTCT  
TTCTAGAAAGGATCACAGCCCCCTGGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGC  
TGCTGTACTGGTTGAATTTGTGTCCCCCTCAAATTCACATCCTTCTTGGAATCTCAGTCTG  
TGAGTTTATTTGGAGATAAGGTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCTATGCTGG  
ATGAAGGTAGACCTAAATTC AATATGACTGGTTTCTTGTATGAAAAGGAGAGGACACAG  
AGACAGAGGAGACGCGGGGAAGACTATGTAAAGATGAAGGCAGAGATCGGAGTTTTCAG  
CCACAAGCTAAGAAACACCAAGGATTGTGGCAACCATCAGAAGCTTGGAAGAGGCAAGA  
AGAATCTTCCCTAGAGGCTTTAGAGGGATAACGGCTCTGCTGAAACCTTAATCTCAGAC  
TTCCAGCCTCCTGAACGAAGAAAGATAAATTTGGCTGTTTTAAGCCACCAAGGATAAT  
TGGTTACAGCAGCTCTAGGAACTAATACAGCTGCTAAAATGATCCCTGTCTCTCGTGT  
TTACATTCTGTGTGTGTCTCCCTCCCACAATGTACCAAAGTTGTCTTGTGACCAATAGAA  
TATGGCAGAAGTGATGGCATGCCACTTCCAAGATTAGGTTATAAAAAGACACTGCAGCTTC  
TACTTGAGCCCTCTCTCTCTGCCACCCACCGCCCCCAATCTATCTTGGCTCAGTCTGCT  
GGGGGAAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAGAGACTTACGTGGTAAAAAAT  
GAAGTCTCCTGCCACAGCCACATTAGTGAACCTAGAGCAGAGACTCTGTGAGATAATC  
GATGTTTGTGTGTTTAAAGTTGCTCAGTTTGGTCTAACTTGTTATGCAGCAATAGATAAA  
TAATATGCAGAGAAAGAG



329/392

## FIGURE 316

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEESVVPNRWLDASLSPVILG  
VQGSQCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWF  
LCTVPEADQPVRLTQLPENGGWNAPIITDFYFQQCD

**N-myristoylation sites:**

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

**Interleukin-1 signature:**

amino acids 111-131

**Interleukin-1 proteins:**

amino acids 8-29, 83-120, 95-134, 64-103

330/392

**FIGURE 317**

ATGGAACTTGGACTTGGAGGCCTCTCCACGCTGTCCCACTGCCCCCTGGCCTAGGCGGCAG  
CCTGCCCTGTGGCCCACCCTGGCCGCTCTGGCTCTGCTGAGCAGCGTCGCAGAGGCCTCC  
CTGGGCTCCGCGCCCCGAGCCCTGCCCCCGCGAAGGCCCCCGCCTGTCTTGGCGTCC  
CCCGCCGGCCACCTGCCGGGGGGACGCACGGCCCGCTGGTGCAGTGGAAGAGCCCGGCGG  
CCGCCCGCGCAGCCTTCTCGGCCCCGCGCCCCGCGCCCTGCACCCCCATCTGCTCTTCCC  
CGCGGGGGCCGCGCGGCGCGGGCTGGGGGCCCGGGCAGCCGCGCTCGGGCAGCGGGGGCG  
CGGGGCTGCCGCTGCGCTCGCAGCTGGTGC CGGTGCGCGCGCTCGGCCTGGGCCACCGC  
TCCGACGAGCTGGTGC GTTTCCGCTTCTGCAGCGGCTCCTGCCCGCGCGCGCTCTCCA  
CACGACCTCAGCCTGGCCAGCCTACTGGGCGCCGGGGCCCTGCGACCGCCCCGGGCTCC  
CGGCCCCGTCAGCCAGCCCTGCTGCCGACCCACGCGCTACGAAGCGGTCTCCTTCATGGAC  
GTCAACAGCACCTGGAGAACCGTGGACCGCCTCTCCGCCACCGCCTGCGGCTGCCTGGGC  
TGA

331/392

**FIGURE 318**

MELGLGGLSTLSHCPWPRRQPALWPTLAALALLSSVAEASLGSA PRSPAPREGPPPPVLAS  
PAGHLPGGRTARWCSEGRARRPPPPQPSRPAPPPPPAPPSALPRGGRAARAGGPGSRARAAGA  
RGCRLRSQ LVPVRALGLGHRSEDLVRFRFCSGSCRRARSPHDLASLLGAGALRPPPGS  
RPVSQPCCRPTRYEAVSFMDVNSTWRTVDRLSATA CGCLG

**signal sequence:**

Amino acids 1-39

**N-glycosylation site:**

Amino acids 202-206

**N-myristoylation sites:**

Amino acids 6-12;67-73;102-108;109-115;119-125

332/392

**FIGURE 319**

GTTGCTATGTTGCCAGGCTGGTCTTGAAGTGCCTTGACCTCCTAAAGTGTGGAACCAC  
AGACGTGAGCCACTCCACCCAGCCTAAACTTCATCTTCTTTGGATGAGATGAACACTTT  
TAACAAGAGAACAGGACTCTATATAAATCGCTGTGGGCTCACCACCTCTAAGGAGGAGCA  
CTGACTGAAGACAGAAAAATTGATGAACTGAAGAAGACATGGTCCATTATGCCTTACAAA  
CTTACACAGTGCTTTGGGAATTCCAAAGTACTCAGTGGAGAGAGGTGTTTCAGGAGCCGT  
AGAGCCAGATCGTCATCATGCTCTGCATTGTGGCTGCTGCTGGGCCTCCTTGCCCTGATGG  
ACTTGTCTGAAAGCAGCAACTGGGGATGCTATGGAAACATCCAAAGCCTGGACACCCCTG  
GAGCATCTTGTGGGATTGGAAGACGTCACGGCCTGAACTACTGTGGAGTTCTGTGCTTCTG  
AAAGCTGGCTGAAATAGACATGCCATACCTCCTGAAATATCAACCCATGATGCAAACCA  
TTGGCCAAAAGTACTGCATGGATCCTGCCGTGATCGCTGGTGTCTTGTCCAGGAAGTCTC  
CCGGTGACAAAATTCTGGTCAACATGGGCGATAGGACTAGCATGGTGCAGGACCCCTGGCT  
CTCAAGCTCCACATCCTGGATTAGTGAGTCTCAGGTTTCCCAGACAACTGAAGTTCTGA  
CTACTAGAATCAAAGAAATCCAGAGGAGGTTTCCAACCTGGACCCCTGACCAGTACCTGA  
GAGGTGGACTCTGTGCCTACAGTGGGGGTGCTGGCTATGTCCGAAGCAGCCAGGACCTGA  
GCTGTGACTTCTGCAATGATGTCCTTGCACGAGCCAAGTACCTCAAGAGACATGGCTTCT  
AACATCTCAGATGAAACCAAGACCATGATCACATATGCAGCCTCAAATGTTACACAGAT  
AAAAGTAGCCAAGGGCACCTGTAAGTGGGAATCTGAGTTTGACCTAAAAGTCATTAAAAT  
AACATGAATCCCATTAATAAAAAAAAAAAAAA

333/392

## FIGURE 320

MSALWLLLGLLLALMDLSESSNWGCYGNIQSLDTPGASCGIGRRHGLNYCGVRASERLAEI  
DMPYLLKYQPMQTIGQKYCMDPAVIAGVLSRKSPGDKILVNMGDRTSMVQDPGSQAPTS  
WISESQVSQTTEVLTTRIKEIQRRFPTWTPDQYLRGGLCAYSGGAGYVRSSQDLSCDFCN  
DVLARAKYLKRHGF

**Important features of the protein:**

**Signal peptide:**

amino acids 1-19

**N-myristoylation sites:**

amino acids 23-29, 26-32, 35-41, 45-51, 50-56, 76-82, 156-162

**Amidation site:**

amino acids 40-44

334/392

**FIGURE 321**

GCCTTATAAAGTAGCCTCTGCATCTGCCTGCCTCGGGCAGAGGAGGGCTACCCCTGGGGCT  
GAGAGTTCACCTGTCTCAGGAACCACCTGAGCCACAGATCCTGTGGGCAGCGGCCAGGG  
CAGCCATGGCTTGGGCAAGTAGGCTGGGCCTGCTGCTGGCACTGCTGCTGCCCGTGGTCG  
GTGCCTCCACGCCAGGCACCGTGGTCCGACTCAACAAGGCAGCATTGAGCTACGTGTCTG  
AAATTGGGAAAGCCCCTCTCCAGCGGGCCCTGCAGGTCACTGTCCCTCATTTCTGGACT  
GGAGTGGAGAGGCGCTTCAGCCCACCAGGATCCGGATTCTGAATGTCCATGTGCCCCGCC  
TCCACCTGAAATTCATTGCTGGTTTCGGAGTGCGCTGCTGGCAGCAGCTAATTTTACTT  
TCAAGGTCTTTTCGCGCCCCAGAGCCCCCTGGAGCTGACGCTGCCTGTGGAAGTCTGGCTG  
ACCCCCGCTGACCCAGAGCTCCATCAGGACCCCTGTGGTCAGCATCTCTGCCTGCTCTT  
TATTCTCGGGCCACGCCAACGAGTTTGATGGCAGTAACAGCACCTCCCACGCGCTGCTGG  
TCCTGGTGACAGCACATTAAAGCTGTCTTGAGTAACAAGCTGTGCCTGAGCATCTCCA  
ACCTGGTGAGGGTGTCATGTCCACCTGGGCACCTTAATTGGCCTCAACCCCGTGGGTCT  
CTGAGTCCCAGATCCGCTATTCCATGGTCAGTGTGCCCACTGTCACCAGTGACTACATTT  
CCCTGGAAGTCAATGCTGTTCTCTTCTGCTGGGCAACCCCATCATCCTGCCCACGGATG  
CCACCCCTTTTGTGTTGCCAAGGCATGTGGGTACCGAGGGCTCCATGGCCACCGTGGGCC  
TCTCCCAGCAGCTGTTTGACTCTGCGCTCCTGCTGCTGCAGAAGGCCGGTGCCCTCAACC  
TGGACATCACAGGGCAGCTGAGGTGCGATGACAACCTGCTGAACACCTCTGCTCTGGGCC  
GGCTCATCCCGGAGGTGGCCCGCCAGTTTCCCGAGCCCATGCCTGTGGTGCTCAAGGTGC  
GGCTGGGTGCCACACCTGTGGCCATGCTCCACACAAACAACGCCACCCTGCGGCTGCAGC  
CCTTCGTGGAGGTCCCTGGCCACAGCCTCCAACCTCGGCTTTCCAGTCCCTCTTCTCCCTGG  
ATGTGGTAGTGAACCTGAGACTCCAGCTCTCTGTGTCCAAGGTGAAGCTTCAGGGGACCA  
CGTCTGTGCTGGGGGATGTCCAGCTCACGGTGGCCTCCTCCAACGTGGGCTTCATTGATA  
CAGATCAGGTGCGCACACTGATGGGCACCGTTTTTGAGAAGCCCCCTGCTGGACCATCTCA  
ATGCTCTCTTGGCCATGGGAATTGCCCTCCCTGGTGTGGTCAACCTCCACTATGTTGCC  
CTGAGATCTTTGTCTATGAGGGCTACGTGGTGATATCCAGTGGACTCTTCTACCAGAGCT  
GAGGCAAGACCACTGGGAGGCCTGAGAGTGGGCCAGCTCGCTGCTCAGGCGAATTTCTCA  
TTTCAAGCCACTGGGGAAACTGAGGCAAAACCATACTTAGTCATCACCAACAAGCTGGAC  
TGCTTAGCTGGGCTGTTTTATCTTCCCTGAGTGCCTGGGTCTCCCTCCCTCACTTCTGCC  
CTTTCCCTTCTCCTCCTCTTCTCCTCCCTCTTCCCTCATCTCCCCCTCCTTCTCTGC  
CCCACCCAGGGGGAGCAGACTGCTCCTCCAGGCTGTATAGACCTGCCCTCTTGCAATTA  
AACAACTTCTCTTGAGCTGC

335/392

**FIGURE 322**

MAWASRLGLLLALLLPVVGASTPGTVVRLNKAALSYVSEIGKAPLQRALQVTVPHFLDWS  
GEALQPTRIRILNVHVPRLHLKFIAGFGVRLAAANFTFKVFRAPEPLELTPVELLADT  
RVTQSSIRTPVVSISACSLFSGHANEFDGSNSTSHALLVLVQKHIKAVLSNKLCLSISNL  
VQGVNVHLGTLIGLNPVGPESQIRYSMVSVPVTSDYISLEVNAVLFLLGNPIILPTDAT  
PFVLPRHVGTEGSMATVGLSQQLFDSALLLLQKAGALNLDITGQLRSDDNLLNTSALGRL  
IPEVARQFPEPMPVVLKVRLGATPVAMLHTNNATLRLQPFVEVLATASNSAFQSLFSLDV  
VVNLRLQLSVSKVKLQGTTSVLGDVQLTVASSNVGFIDTDQVRTLMGTVFEKPLLDHLNA  
LLAMGIALPGVVNLHYVAPEIFVYEGYVVISSGLFYQS

**Important features of the protein:****Signal peptide:**

Amino acids 1-20

**Transmembrane domain:**

Amino acids 217-236

**N-glycosylation sites:**

Amino acids 96-100;151-155;293-297;332-336

**N-myristoylation sites:**

Amino acids 8-14;149-155;189-195;249-255;252-258;283-289

**LBP / BPI / CETP family proteins:**

Amino acids 22-50; 251-287

336/392

**FIGURE 323**

TTGAAAATCTACTCTATCAGCTGCTGTGGTTGCCACCATTCTCAGGACCCTCGCCATGAA  
AGCCCTTATGCTGCTCACCTGTCTGTTCTGCTCTGCTGGGTCTCAGCTGACATTGCTG  
TCACTCCTGCTACAAGGTCCCTGTGCTGGGCTGTGTGGACCGGCAGTCCTGCCGCCTGGA  
GCCAGGACAGCAATGCCTGACAACACATGCATACCTTGGTAAGATGTGGGTTTTCTCCAA  
TCTGCGCTGTGGCACACCAGAAGAGCCCTGTCAGGAGGCCTTCAACCAAACCAACCGCAA  
GCTGGGTCTGACATATAACACCACCTGCTGCAACAAGGACAACTGCAACAGCGCAGGACC  
CCGGCCCACTCCAGCCCTGGGCCTTGTCTTCCTTACCTCCTTGGCTGGCCTTGGCCTCTG  
GCTGCTGCACTGAGACTCATTCCATTGGCTGCCCCCTCCTCCACCTGCCTTGGCCTGAGC  
CTCTCTCCCTGTGTCTCTGTATCCCCTGGCTTTACAGAATCGTCTCTCCCTAGCTCCCAT  
TTCTTTAATTAAACACTGTTCCGAGTGGTCTCCTCATCCATCCTTCCCACCTCACACCCT  
TCACTCTCCTTTTTCTGGGTCCCTTCCCACCTTCTTCCAGGACCTCCATTGGCTCCTAGA  
AGGGCTCCCCACTTTGCTTCCCTATACTCTGCTGTCCCCTACTTGAGGAGGGATTGGGATC  
TGGGCCTGAAATGGGGCTTCTGTGTTGTCCCCAGTGAAGGCTCCCACAAGGACCTGATGA  
CCTCACTGTACAGAGCTGACTCCCCAAACCCAGGCTCCCATATGTACCCCATCCCCATA  
CTCACCTCTTTCCATTTTGAGTAATAAATGTCTGAGTCTGGAAAAAAAAAAAAAAAAAAAA



337/392

## FIGURE 324

MKALMLLTLSVLLCWVSADIRCHSCYKVPVLGCVDROSCRLEPGQQCLTTHAYLGKMWVF  
SNLRCGTPEEPCQEAFNQTNRKLGTYNTTCCNKDNCNSAGPRPTPALGLVFLTSLAGLG  
LWLLH

**Important features of the protein:**

**Signal peptide:**

amino acids 1-18

**N-glycosylation sites:**

amino acids 77-81, 88-92

**N-myristoylation site:**

amino acids 84-90

**Ly-6 / u-PAR domain protein signature:**

amino acids 85-98

338/392

**FIGURE 325**

ACGGGCCG CAGCGGCAGTGACGTAGGGTTGGCGCAGCGATCCGTTGCGGCTGCAGCTCTG  
CAGTCGGGCCGTTCCCTTCGCCGCCGCCAGGGGTAGCGGTGTAGCTGCGCAGCGTCGCGCG  
CGCTACCGCACCCAGGTTCCGGCCCGTAGGCGTCTGGCAGCCCGGCCGCGCCATCTTCATCGAG  
CGCCATGGCCG CAGCCTGCGGGCCGGGAGCGGCCGGGTACTGCTTGCTCCTCGGCTTGCA  
TTTGTTTCTGCTGACCGCGGGCCCTGCCCTGGGCTGGAACGACCCTGACAGAATGTTGCT  
GCGGGATGTAAAAGCTCTTACCCTCCACTATGACCGCTATACCACCTCCCGCAGGCTGGA  
TCCCATCC CACAGTTGAAATGTGTTGGAGGCACAGCTGGTGTGATTCTTATACCCAAA  
AGTCATACAGTGT CAGAACAAGGCTGGGATGGGTATGATGTACAGTGGGAATGTAAGAC  
GGACTTAGATATTGCATACAAATTTGGAAAACTGTGGTGAGCTGTGAAGGCTATGAGTC  
CTCTGAAGACCAGTATGTACTAAGAGGTTCTTGTGGCTTGGAGTATAATTTAGATTATAC  
AGAACTTGGCCTGCAGAACTGAAGGAGTCTGGAAAGCAGCACGGCTTTGCCTCTTCTC  
TGATTATTATTATAAGTGGTCCTCGGCGGATTCTGTAAACATGAGTGGATTGATTACCAT  
CGTGGTACTCCTTGGGATCGCCTTTGTAGTCTATAAGCTGTTCTGAGTGACGGGCAGTA  
TTCTCCTCCACCGTACTCTGAGTATCCTCCATTTTCCACCGTTACCAGAGATTCACCAA  
CTCAGCAGGACCTCCTCCCCCAGGCTTTAAGTCTGAGTTCACAGGACCACAGAATACTGG  
CCATGGTGCAACTTCTGGTTTTGGCAGTGCTTTTACAGGACAACAAGGATATGAAAATTC  
AGGACCAGGGTTCTGGACAGGCTTGGGAACTGGTGAATACTAGGATATTTGTTTGGCAG  
CAATAGAGCGGCAACACCCTTCTCAGACTCGTGGTACTACCCGTCCTATCCTCCCTCCTA  
CCCTGGCACGTGGAATAGGGCTTACTCACCCCTTCATGGAGGCTCGGGCAGCTATTCGGT  
ATGTTCAAACCTCAGACACGAAAACCAGAACTGCATCAGGATATGGTGGTACCAGGAGACG  
ATAAAGTAGAAAGTTGGAGTCAAACACTGGATGCAGAAATTTTGGATTTTTCATCACTTT  
CTCTTTAGAAAAAAGTACTACCTGTTAACAATTGGGAAAAGGGGATATTCAAAGTTCT  
GTGGTGTTATGTCCAGTGTAGCTTTTGTATTCTATTATTTGAGGCTAAAAGTTGATGTG  
TGACAAAATACTTATGTGTTGTATGTCAGTGTAAACATGCAGATGTATATTGCAGTTTTTG  
AAAGTGATCATTACTGTGGAATGCTAAAAATACATTAATTTCTAAACCTGTGATGCCCT  
AAGAAGCATTAAGAATGAAGGTGTTGTACTAATAGAACTAAGTACAGAAAATTTTCAGTT  
TTAGGTGGTTGTAGCTGATGAGTTATTACCTCATAGAGACTATAATATTCTATTTGGTAT  
TATATTATTTGATGTTTGCTGTTCTTCAAACATTTAAATCAAGCTTTGGACTAATTATGC  
TAATTTGTGAGTTCTGATCACTTTTGGAGCTCTGAAGCTTTGAATCATTCAGTGGTGGAGA  
TGGCCTTCTGGTAACTGAATATTACCTTCTGTAGGAAAAGGTGGAAAATAAGCATCTAGA  
AGGTTGTGTGAATGACTCTGTGCTGGCAAAAATGCTTGAAACCTCTATATTTCTTTCGT  
TCATAAGAGGTAAAGGTCAAATTTTCAACAAAAGTCTTTTAATAACAAAAGCATGCAGT  
TCTCTGTGAAATCTCAAATATTGTTGTAATAGTCTGTTTCAATCTTAAAAAGAATCA

339/392

**FIGURE 326**

MAAACGPGAAGYCLLLGLHLFLLTAGPALGWNDPDRMLLRDVKALTLHYDRYTTSRRLLDP  
IPQLKCVGGTAGCDSYTPKVIQCQNKGDGYDVQWECKTDLDIAYKFGKTVVSCEGYESS  
EDQYVLRGSCGLEYNLDYTELGLOKLKESGKQHGFASFSDYYYKWSSADSCNMSGITIV  
VLLGIAFVVYKLFLSDGQYSPPPYSEYPPFSHRYQRFTNSAGPPPPGFKSEFTGPQNTGH  
GATSGFGSAFTGQQGYENSGPGFWTGLGTGGILGYLFGSNRAATPFSDSWYYPSYPPSY  
GTWNRAYSPLHGGSGSYSVCSNSDTKTRTASGYGGTTRR

**Signal peptide:**  
amino acids 1-30

**Transmembrane domain:**  
amino acids 171-190

**N-glycosylation site:**  
amino acids 172-176

**Glycosaminoglycan attachment sites:**  
amino acids 244-248, 259-263, 331-335

**Tyrosine kinase phosphorylation site:**  
amino acids 98-106

**N-myristoylation sites:**  
amino acids 68-74, 69-75, 131-137, 241-247, 247-253, 266-272,  
270-276, 278-284, 312-318

340/392

**FIGURE 327**

GGCACGAGGTGGAAGGGCTTTTACAAACAGATTGCTGGCCCCACCCCCCAGAATTTCTCA  
TCAGGAGTGGGCAAGACCAATCATTTGCATTTCTGACAAGTTCCCAGGAGCTGCAGCTGC  
TGGCCCTGGAACCACACTTTGAGAACCACTGCTTTAGACCAAACACCAAAGGAAGATGCA  
GCCACCCTCCTTTACATGTCACAACGCTCAGGGTCCATGAGTACCTCAGGCTGTCCAGCT  
GAGCTCCACCTGCAGCAGCCGAGATTCCCGACTCGCTCCACCATTGGGGGCTAGGAGTGA  
AGCGTGTCAACATGGTCAGCTCATGGCCAGCCAGGAAAGCCTCTCTGCTGTGCGTCTGTG  
CAGTTCTTGTTCTTCCCTGGAGGACTCTTGATCGCCTGTGATCTTGGCCAGGAGACCAG  
GTGCCTGGGTCCCTTCCTGGAAGGGGACAAGTTACACACCCAGCCCCATTTCCACCA  
ACTTCTACATGCCTTGGGAGAACCTTCTACATGTTGGCTGCCCCCTTCCCTATTTTCAGC  
AGTGCCAGTCCTGCTTATAAACCTGAGGCCTGCTCCCCATACCTTCCCTGTGCAAGTGC  
CAGCCGTTATTCAGGCAGCCCAATGTTGTTGAGGCCAGATGGATTCCTGGAAGCAGCTG  
GCCATGGATGTGAGTCATCACAGTATTCTAGAAACAGAGAAGAGGTCTTAACCTAATGC  
GCATAGAGAAATTGTTCTCATTGTAAACATAACCCCTGTCCTTAGCTGATCTAGGTGGAAG  
CCCAGCTTCATGTGCTAGGGGGCATGATAATGATAATAAGGAATTGTATCTAGGACTAA

341/392

## FIGURE 328

MVSSWPARKASLLCVC AVLVL PWRTL GSPVILARRPGAWVPSWKGTSYTPQPHFPTNFYM  
PWENLLHVGCP LPLFQQCPVLLINLRPAPHTFPVQVPAVIPGSPMLLRPDGFLEAAGPWM

Signal peptide:  
amino acids 1-27

cAMP- and cGMP-dependent protein kinase phosphorylation site:  
amino acids 8-12

342/392

**FIGURE 329**

CAAAGAGTAGTCAGTCCCTTCTTGGCTCTGCTGACACTCGAGCCCACATTCCATCACCTG  
CTCCCAATCATGCAGGTCTCCACTGCTGCCCTTGCCGTCCTCCTCTGCACCATGGCTCTC  
TGCAACCAGGTCTCTCTGCACCACTTGCTGCTGACACGCCGACCGCCTGCTGCTTCAGC  
TACACCTCCCGACAGATTCCACAGAATTTCATAGCTGACTACTTTGAGACGAGCAGCCAG  
TGCTCCAAGCCCAGTGTCATCTTCCTAACCAAGAGAGGCCGGCAGGTCTGTGCTGACCCC  
AGTGAGGAGTGGGTCCAGAAATACGTCAGTGACCTGGAGCTGAGTGCCGTGAGGGGTCCAG  
AAGCTTCGAGGCCCAGCGACCTCAGTGGGCCCAGTGGGGAGGAGCAGGAGCCTGAGCCTT  
GGGAACATGCGTGTGACCTCTACAGCTACCTCTTCTATGGACTGGTTATTGCCAAACAGC  
CACACTGTGGGACTCTTCTTAACTTAAATTTTAAATTTATTTATACTATTTAGTTTTTATA  
ATTTATTTTTTGATTTACAGTGTGTTTGTGATTGTTTGCTCTGAGAGTTCCCCCTGTCCC  
CTCCCCCTTCCCTCACAGTGTGTCTGGTGACAACCGAGTGGCTGTCATCGGCCTGTGTAG  
GCAGTCATGGCACCAAAGCCACCAGACTGACAAATGTGTATCAAATGCTTTTGTTCAGGG  
CTGTGATCGGCCTGGGGAAATAATAAGATGTTCTTTTAAACGGTAAAAAA

343/392

## FIGURE 330

MQVSTAALAVLLCTMALCNQVLSAPLAADTPTACCFSYTSRQIPQNFIADYFETSSQCSK  
PSVIFLTRGRQVCADPSEEWVQKYVSDLELSA

Signal sequence:

1-23

Small cytokines (intercrine/chemokine) C-C subfamily  
signature:

1-35, 2-36, 10-44, 34-74, 50-90

Small cytokines (intecrine/chemokine):

24-89

344/392

**FIGURE 331**

GGCACGAGGTGAGACTTTAAATGAAATGTCTCACAAGCTAGGTGATCCAGGTTTTGTGGT  
CTTTGCAACCCTTGTGGTCATTGTGGCCTTGATATTAATCTTCGTGGTGGGTCCCTCGCCA  
TGGACAGACAAACATTCTTGTGTACATAACAATCTGCTCTGTAATCGGCGCGTTTTTCAGT  
CTCCTGTGTGAAGGGCCTGGGCATTGCTATCAAGGAGCTGTTTGCAGGGAAGCCTGTGCT  
GCGGCATCCCCTGGCTTGGATTCTGCTGCTGAGCCTCATCGTCTGTGTGAGCACACAGAT  
TAATTACCTAAATAGGGCCCTGGATATATTCAACACTTCCATTGTGACTCCAATATATTA  
TGTATTCTTTACAACATCAGTTTTAACTTGTTTCACTATTCTTTTTAAGGAGTGGCAAGA  
TATGCCTGTTGACGATGTCATTGGTACTTTGAGTGGCTTCTTTACAATCATTGTGGGGAT  
ATTCTTGTTCATGCCTTTTAAAGACGTCAGCTTTAGTCTAGCAAGTCTGCCTGTGTCTTT  
TCGAAAAGACGAGAAAGCAATGAATGGCAATCTCTCTAATATGTATGAAGTCTTAATAA  
TAATGAAGAAAGCTTAACCTGTGGAATCGAACAACACACTGGTGAAAATGTCTCCCGAAG  
AAATGGAAATCTGACAGCTTTTTAAAGAAAGGTGTAATTAAAGGTAAATCTGTGATTGTTA  
TGAAGTGAATTTGAATATCATCAGAATGTGTCTGAAAAAACATTGTCCTCAAATAATGTT  
CTTTAAAGGCAATCTTTTTAAAGATTTCACTAATTTGGACCAAGAAATTACTTTTCTTGT  
ATTTAAACAAACAATGGTAGCTCACTAAAATGACCTCAGCACATGACGATTTCTATTAAC  
ATTTTATTGTTGTAGAAGTATTTTACATTTTCATCCCTTCTCCAAAAGCCGAATGACTA  
ATGACAGTTTTAAGTCTATGAAAATGCTTTATTTTTTTCATTGGTGATGAAAGTCTGAAAT  
GTGCATTTGTTCATCCCCACTCCATCAATCCCTGACCATGTAAGGCTTTTTTATTTTAAAA  
AAACAGAGTTATCCCAATACATTATCCTGTGATTTACCTTACCTACAAAAGTGGCTCCTG  
TTTGTGTTGATGATGATTGGTTTTATTTTTGAAATATTTATTAAGGGAAAATAAGTTACT  
GAATGAAGGAACCTCTTTCTTACAAAACAAAAAAGGGCAGAAATCACCCCAAGGAACG  
ATTTCTCAGGTTGAGATGATCACCGTGAATCCGGCTTCCTCTGAGCATTGATGGCCTTA  
GCACCTCATCAAGCCAGCACATCCTGCCTGCTGTTGCAGCCTGGCTGGGTTTATTCTTCA  
GTTACCTAATCCCATGATGCCTGGAACCTTGATTACCGTTTTTACATCAGCTCTTGTACT  
TTTCAGTATATTTTTCATAATGAGTTATATTGTCAATTTAGACTTTGAACAGCTCTGGGAAA  
TAGAAGACTAGGGTTGTTTCTTAAATTTAGCTCATGTTATAATAAAAAGTTGAAATG



345/392

## FIGURE 332

MSHKLGDPGFVVFATLVVIVALILIFVVGPRHGQTNILVYITICSVIGAFSVSCVKGLGI  
AIKELFAGKPVLRHPLAWILLLSLIVCVSTQINYLNRLDIFNTSIVTPIYYVFFTTSVL  
TCSAILFKEWQDMPVDDVIGTLSGFFTIIVGIFLLHAFKDVSFSLASLPVSFRKDEKAMN  
GNLSNMYEVLNNNEESLTCGIEQHTGENVSRRNGNLTA

Signal sequence:

1-33

Transmembrane domain:

40-60, 70-90, 103-123, 139-159

N-glycosylation site:

103-106, 182-185, 208-211, 215-218

N-myristoylation site:

57-62, 140-145, 181-186, 214-219

346/392

**FIGURE 333**

GTGATGGCGGCTGGTGATGGGGACCTGAAGCTAGGCACCCTGGGGAGTGGCAGCGAGAGC  
AGCAACGACGGCGGCAGCGAGAGTCCAGGCGACGCGGGAGCGGCAGCGGAAGGGGGAGGC  
TGGGCGGCGGCGGCGTTGGCGCTTCTGA<sup>c</sup>GGGGGGCGGGGAAATGCTGCTGAACGTGGCG  
CTGGTGGCTCTGGTGCTGCTGGGGGCC<sup>t</sup>TACCGGCTGTGGGTGCGCTGGGGGCGGCGGGGT  
CTGGGGGCCGGGGCCGGGGCGGGCGAGGAGAGCCCCGCCACCTCTCTGCCTCGCATGAAG  
AAGCGGGA<sup>c</sup>CTTCAGCTTGGAGCAGCTGCGCCAGTACGACGGCTCCCGCAACCCGCGCATC  
CTGCTCGCGGTCAATGGGAAAGTCTTCGACGTGACCAAAGGCAGCAAGTTCTACGGCCCCG  
GCGGGTCCATATGGAATATTTGCTGGTAGGGATGCCTCCAGAGGACTGGCCACATTTTGC  
CTAGATAAAGATGCACTTAGAGATGAATATGATGATCTCTCAGATTTGAATGCAGTACAA  
ATGGAGAGTGTTGAGAATGGGAAATGCAGTTTAAAGAAAAATATGATTATGTAGGCAGA  
CTCCTAAAACCAGGAGAAGAACCATCAGAATATACAGATGAAGAAGATACCAAGGATCAC  
AATAAACAGGATTGAACTTTGTAAACAACCAAAGTCAGGGGCCTTCAGAACTGCAATTCT  
TACTCCCTTTCACAGACTGTCCGGAGTCTTTGGGTTTGATTACCTGCTGCGAAAAACAT  
TCAACAAATTGTGTACAAGATAAATTAATCTCACTATGAAGATTTGAATAACTAGACATT  
ATTTATGCTGCCAAACTCATTGTGTCAGTTGTTTGTAAATGTCTAGTGGGGCTTCATCAT  
CCTGAAAAGAAGGAGACAGGGATTTTTTTAAAGAGCAAGAAAGTCACAAATATTACTTCTT  
TCCTTCCTTTTTTCCTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTTAAATATATTG  
AAGACAACCAGATATGTATTTGCTACTCAAGTGTACAGATCTCCTCAAGAAACATCAAGG  
G

347/392

## FIGURE 334

MAAGDGDVKLGTLGSGSESSNDGGSESPGDAGAAAEGGGWAAAALALLTGGGEMLLNVAL  
VALVLLGAYRLWVRWGRRGLGAGAGAGEESPATSLPRMKKRDFSLEQLRQYDGSRNPRIL  
LAVNGKVFDTVTKGSKFYGPAGPYGIFAGRDASRGLATFCLDKDALRDEYDDLSDLNAVQM  
ESVREWEMQFKEKYDYVGRLLKPGEEPSEYTDDEEDTKDHNKQD

Signal sequence:

None

Transmembrane domain:

45-65

Tyrosine kinase phosphorylation site:

202-210

N-myristoylation site:

11-16, 16-21, 37-42, 38-43, 79-84, 81-86, 83-88, 144-149

Amidation site:

75-78



349/392

**FIGURE 336**

TGRGYCGDHESSEFGAMEEPGATPQPYLGLLLEELRRVVAALPEGMRPDSNLYGFPWELVI  
CAAVVGFFAVLFFLWRSFRSRSRLYVGREKKLALMLSGLIIEKSKLLEKFSLVQKEYEG  
YEVESLKDASFEKEATEAQSLEATCEKLNRSNSELEDEILCLEKELKEEKSKHSEQDEL  
MADISKRIQSLEDESKSLKSQVAEAKMTFQIFQMNEERLKIAIKDALNENSQLOESQKQL  
LQEAENVWKEQVSELNKQKVTTFEDSKVHAEQVLNDKESHIKTLTERLLKMKDWAAMLGEDI  
TDDDNLELEMNSESENGAYLDNPPKGALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDK  
TKEELTEHIKNLQTQQASLQSENTHFENENQKLQQKLKVMTELYQENEMKLHRKLTVEEN  
YRLEKEEKLSKVDEKISHATEELETYRKRAKDLLEELERTIHSYQGQIISHEKKAHDNWL  
AARNAERNLNDLRKENAHNRQKLTETELKFELLEKDPYALDVPNTAFGRGSRGPGNPLDH  
QITNERGESSCDRLTDPHRAPSDTGSLSPWDQDRRMMFPPPGQSYPDALPPQRQDRFC  
SNSGRLSGPAELRSFNMPSLDKMDGSMPSSESSRNDTKDDLGNLNVDPSSLPAENEATG  
PGFVPPPLAPIRGPLFPVDARGPFLRRGPPFPPPPPGAMFGASRDYFPPRDFPGPPPAPF  
AMRNVYPPRGFPFPPYLPPRPGFFPPPPHSEGRSEFFPSGLIPPSNEPATEHPEPQQET

**Signal sequence:**

None

**Transmembrane domain:**

54-74

**N-glycosylation site:**

150-153, 338-341, 636-639

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

413-416

**Tyrosine kinase phosphorylation site:**

414-421

**N-myristoylation site:**

466-417, 625-630, 697-702

**Leucine zipper pattern:**

142-163

350/392

**FIGURE 337**

GGACTGCGGTCTCGGGCAGCAATGGCCGAGAAGCGCGACACACGGGACTCCGAAGCCCAG  
CGGCTCCCCGACTCCTTCAAGGACAGCCCCAGTAAGGGCCTTGGACCTTGC GGATGGATT  
TTGGTGGCGTTCTCATTCTTATTCACCGTTATAACTTTCCCAATCTCAATATGGATGTGC  
ATAAAGATTATAAAAGAGTATGAAAGAGCCATCATCTTTAGATTGGGTGCGATTTTACAA  
GGAGGAGCCAAAGGACCTGGT TTTGTTTTTATTCTGCCATGCACTGACAGCTTCATCAAA  
GTGGACATGAGAACTATTTCA TTTGATATTCTCCTCAGGAGATCCTGACAAAGGATTCA  
GTGACAATTAGCGTGGATGGTGTGGTCTATTACCGCGTTCAGAATGCAACCCCTGGCTGTG  
GCAAATATCACCAACGCTGACTCAGCAACCCGTCTTTTGGCACAACTACTCTGAGGAAT  
GTTCTGGGCACCAAGAATCTTTCTCAGATCCTCTCTGACAGAGAAGAAATTGCACACAAC  
ATGCAGTCTACTCTGGATGATGCCACTGATGCCTGGGGAATAAAGGTGGAGCGTGTGGAA  
ATTAAGGATGTGAACTACCTGTGCAGCTCCAGAGAGCTATGGCTGCAGAACGAGCG  
TCCCGCGAGGCCCGCGCCAAGGTTATTGCAGCCGAAGGAGAAATGAATGCATCCAGGGCT  
CTGAAAGAAGCCTCCATGGTCACTGAATCTCCTGCAGCCCTTCAGCTCCGATACCTG  
CAGACACTGACCACCATTGCTGCTGAGAAAACTCAACAATTGTCTTCCCTCTGCCATA  
GATATGCTGCAAGGAATCATAGGGGCAAAACACAGCCATCTAGGCTAGTGTAGAGATGAG  
CGCTAGCCTTCCAAGCATGAAGTCGGGGACCAAATTAGCCTTTAACTCATAAAGAGAGGG  
TAGGGCTTTTCTTTTCCATATGTCAATTGTGGTGTTC CAGAATGTATAGCAGTTATAA  
AAATAGGTGAAAGAATTGTTAGCTTGTAATACTGAGAGATTGGTGATTTATATAAGGTA  
ATCTGTTAGTCTTAAATAGTTAAAAGTTTGTATTTT TAGATTATTATGTAGTAGGTTAG  
ATCCCTCTTGTTT GACTTCCACTGACTCATTCTGAACCCCTAAGCACCCAGGCCACAG  
GCAAGAACCTGGGCTGTA ACTGCCACCTGACACCGCTGACTGGCTAAATGCTTTGCAGAA  
AGTGATGACCTTACACCACAACCAGCTTCTCCAGGTCATATGTGCCTTACCTCCAGAAGT  
CTTTTTTTTTTTTTTTTTCTGAGATGGAGTTTCACTCTTGTTGCCAGGCTGGAGTGCAA  
TAGCATGATCTCGGCTCACTGCAACCTCCGCTCCTGGGTTC AAGAGATTCTCCTGCCTC  
AGCCTCCCCAGTAGCTGGGATTACAGGCTCATGCCACCATGCCCAGCTAATTTTTGTATT  
ATTATTATTGTTTTTTAGTAGAGACGGGTTTCACCATGTTGGCCAGGCTAGTCACGAAC  
TCCTAACCTCAGGTGATCCACCCACCTCTGCCTCCAAAGTGCTGGATTACAGGCTGAGCT  
ACCACCTGGT TTTGGAGAGTCTTAATTAATTGAAATTTCCCTAATGTTTCATTTATTTTCT  
AAATCCAGCCGTGTTTCAGAATAATCCTTACTTGAGAGTAGCCATTTTCTTGTTACTTG  
TCAGAACTAGAGGAAATAGCCAAGACTAATGAAAAACATTACTCTAACCTTAAAGACT  
TTTAAATTCACTACTAGAGTGGTCATTTTAAAATACATCCATGTTTTAACTTATTTTGA  
GCCTTCTTTTATGAGTAAATGATTCCTCCTTGTTCTGTCTTTCAAACCAGCTAAATATT  
TGTGACAAAAGTGACTTTTTTCTCACTGTGCCTATTTTCATATATCAGGTTTTAAATAG  
TTTTTAATTTTTTAATAAAATTTTTCTCTACGTTCTATATGCAATTGTTATATATCTATTT  
GAATAGCTGAAGGACTAAAATACTTTTTTAAGAGATAA CTTCAGGAAACCATTATATTTT  
ACTATCTGCATGCTGT TAACTGTGGTACACTGTGAAATATGTTGATTACAAACCCATTCA  
TTACATAGTATAAGGAATTCACAGTATATTGACTATATAGTGTCTAATGACTGGGCAGAT  
ACTGTCAACTTACAATATCTATATAGAGAGGCTTTAAACTTACCTTACTCATTTCTCTATG  
ATGTATGACTTGATGCTGAAAGAGGAAGCTGGTCAGCTCCTCATGGACAACAAATTCCTTA  
GTCTATAATATTAGGAGACATCTCTAGTTTTTGCAAATGTCTGTGAATCTGAGCAACCTGG  
ACTTCTGCTTACTGGCCAGAAAGCTGGCGGGTGACATTTGTAACATTTCTCTTTGAGAC  
TCTGAGTTCACCTAGAGAAGTCTAAGCATAACAGCTTTCTTTCCAGCACGAGCCTTTAT  
AGCTCTCTTTAGCTCAACCACTCTGTCCATCCAGCCAATGGATGTCCTTCCCTGTACCCA  
ATTCAAGCTTATTTTAGGGAAGCCTTGAAACTACCATGTATCTGGCTCTAGCTGAGTTAT  
TGAGGATTGAGCCAGTGCAACGTTAAACTCAGTGCACCTTACATTTGATTTAAATGATGGT  
TTTATCTGTTGTGTGAAGTGGTTCACCCCTGAGGACCAGGAGCCTCCATATCCTGACTGA  
AAACCTTTTCTGAGACTTAGAGTAACAGTACTTTTGGTTCTTGAGTTCTCCTGTCTCCA  
GATACCTAAATGACCTTGACTTTTTCTGCCTTGTGAATTCGTAGTCCAATCAGCTGAAATT  
AAATCACTTGGGAGGGACGCATAGAAGGAGCTCTAGGAACACAGTGCCAGTGCAGAAAGTT  
TCTCCAGGTGGCCTCCCTTTCCAACAATGTACATAATAAAGTGTATGCACCTTCACT

351/392

**FIGURE 338**

MAEKRDTRDSEAQRLPDSFKDSPSKGLGPCGWILVAFSFLFTVITFPISIWMCIKIKEY  
ERAIIFRLGRILQGGAKGPGLFFILPCTDSFIKVD MRTISFDIPPQEILTKDSVTISVDG  
VYYRVQNATLAVANITNADSATRLLAQTTLRNVLGTKNLSQILSDREEIAHNMQSTLDD  
ATDAWGIKVERVEIKDVKLPVQLQRAMAAEAEASREARAKVIAAEGEMNASRALKEASMV  
ITESPAALQLRYLQTLTTIAAEKNSTIVFPLPIDMLQGIIGAKHSHLG

Signal sequence:

1-45

Transmembrane domain:

None

N-glycosylation site:

128-131, 135-138, 159-162, 229-232, 264-267

cAMP- and cGMP-dependent protein kinase phosphorylation  
site:

4-7

N-myristoylation site:

26-31, 278-283, 281-286

SPFH domain/Band 7 family:

39-230

352/392

**FIGURE 339**

TCTAGAGCCCTCTCCCAACATGGCGGCCTCAGCAAAAAAGAAGAATAAGAAGGGGAAGAC  
TATCTCCCTAACAGACTTTCTGGCTGAGGATGGGGGTACTGGTGGAGGAAGCACCTATGT  
TTCCAAACCAGTCAGCTGGGCTGATGAAACGGATGACCTGGAAGGAGATGTTTCGACCAC  
TTGGCACAGTAACGATGACGATGTGTATAGGGCGCCTCCAATTGACCGTTCCATCCTTCC  
CACTGCTCCACGGGCTGCTCGGGAACCCAATATCGACCGGAGCCGTCTTCCCAATCGCC  
ACCCTACACTGCTTTTCTAGGAAACCTACCCTATGATGTTACAGAAGAGTCAATTAAGGA  
ATTCTTTCGAGGATTAAATATCAGTGCAGTGCCTTTACCACGTGAACCCAGCAATCCAGA  
GAGGTTGAAAGGTTTTGGTTATGCTGAATTTGAGGACCTGGATTCCCTGCTCAGTGCCCT  
GAGTCTCAATGAAGAGTCTCTAGGTAACAGGAGAATTGAGTGGACGTTGCTGATCAAGC  
ACAGGATAAAGACAGGGATGATCGTTCTTTTGGCCGTGATAGAAATCGGGATTCTGACAA  
AACAGATACAGACTGGAGGGCTCGTCCTGCTACAGACAGCTTTGATGACTACCCACCTAG  
AAGAGGTGATGATAGCTTTGGAGACAAGTATCGAGATCGTTATGATTTCAGACCGGTATCG  
GGATGGGTATCGGGATGGGTATCGGGATGGCCACGCCGGGATATGGATCGATATGGTGG  
CCGGGATCGCTATGATGACCGAGGCAGCAGAGACTATGATAGAGGCTATGATTCCCGGAT  
AGGCAGTGGCAGAAGAGCATTTGGCAGTGGGTATCGCAGGGATGATGACTACAGAGGAGG  
CGGGGACCGCTATGAAGACCGATATGACAGACGGGATGATCGGTCGTGGAGCTCCAGAGA  
TGATTACTCTCGGGATGATTATAGGCGTGATGATAGAGGTCCCCCAGAACCCAACT  
GAATCTAAAGCCTCGGAGTACTCCTGAAGAAGATGATTCTCTGCTAGTACCTCCAGTC  
CACTCGAGCTGCTTCTATCTTTGGAGGGGCAAAGCCTGTTGACACAGCTGCTAGAGAAAG  
AGAAGTAGAAGAACGGCTACAGAAGGAACAAGAGAAGTTGCAGCGTCAGTGAATGAGCC  
AAAAGTAGAACGACGGCCTCGGGAGAGACACCCAAGCTGGCGAAGTGAAGAACTCAGGA  
ACGGGAACGGTTCGAGGACAGGAAGTGAGTCATCACAACCTGGGACCTCCACCACATCTAG  
CAGAAATGCACGAAGGAGAGAGAGTGAAGTCTCTAGAAAATGAAACACTCAATAAGGA  
GGAAGATTGCCACTCTCCAACCTTCTAAACCTCCCAAACCTGATCAGCCCCCTAAAGGTAAT  
GCCAGCCCCCTCCACCAAAGGAGAATGCTTGGGTGAAGCGAAGTTCTAACCTCCTGCTCG  
ATCTCAGAGCTCAGACACAGAGCAGCAGTCCCCCTACAAGTGGTGGGGGAAAAGTAGCTCC  
AGCTCAACCATCTGAGGAAGGACCAGGAAGGAAAGATGAAAATAAAGTAGATGGGATGAA  
TGCCCCAAAAGGCCAACTGGGAACTCTAGCCGTGGTCCAGGAGACGGAGGGAAACAGAGA  
CCACTGGAAGGAGTCAGATAGGAAAGATGGCAAAAAGGATCAAGACTCCAGATCTGCACC  
TGAGCCAAAGAAACCTGAGGAAAATCCAGCTTCTAAGTTCAGTCTGCAAGCAAGTATGC  
TGCTCTCTCTGTTGATGGTGAAGATGAAAATGAGGGAGAAGATTATGCCGAATAGACCTC  
TACATCCTGTGCTTTTCTCCTAGTTTCTCTCCACCCTGGAAACATTCGAGAGCAAAATCAA  
ACCTCTATCCAGACAAGACAAAATAAACTCAACATCTCCTGAAGACCTTTCTTACCTTT  
TTTTAAAAACAAAAXTGAAATTATTTTGCATGCTGCTGCAGCCTTTAAAGTATTGAAGT  
AACTGGAGAATTGCCAATACAGCCAGAGAGAAAAGGACTACAGCTTTTTTAGAGGAAAAGT  
TGTGGTGCGTTATGTCACCATGCAGTTGCCAGTGTGATTAGTGCCTAGGGGTCTCATTTA  
GCAGAAATGGTAATGACAGTGTATATAATGCCTGGAACCTGGTTGGGCAGTAGGGGAGGGA  
GGTAGAAGGAAAAGTGTGAGATTTCTACCTTTTAGTTTTATCCTATTGTGGCATATATG  
AATTCTCAAACATTATCTGAATAAATTTTCCACTCTTGGAAGGTAGATTTAGCCTCAAG  
TTGTTCTAGTCTCCAGGAGGCTGCCAGCCCCCTCCTCTTATTTAATTCTGAGTTTGGGGG  
CCAGCCTAGAGGGAATTCCTTTTTTTTTTTTTTTTAAACCCCCAGGGGGTAGTTGGGAGT  
GAGACTATAGGCCATAAAGAATGGGACTGCATTGGACCAAAATAAATGGGAAAATCGTGG  
TTTGAAAAGAAGCTTTTGGGAAGTGATGAGTCATTTTGCACCAGGTAATAGGGGAAAATT  
GTGTGACCTCCAGCAAACACATGAATGGTTATTTCTTGGAGCCGGAAGCACTTGGGGGTC  
GTGGTAATTCCAGTGTTTTCTGTGTCTAGTTTACCCTTTCTAAACACTGTCTTTTTT  
GAAAGTTTTGAATATATCCACATTTCTATTGAAACCTTGAACTAAAAATTTAGACTCTTA  
TCGTCACTCTTAAGTTCTTCATGCTACTCTTAACCTCCCAAAAAGCAGTATCTAAGTCACA  
TACATGATGCTTGGGCATTTTCTGAGCCATGGAGAACTCTGAAAGGAAGAATCGCTGCT  
TTTCTCAAGCAAATCGGTTTCTTGATGTCTTTTGGTTCTCCTTGCCTGCTCCTGATGCTT



353/392

GGACCCCTTTTATTGATCAGAGTGCTCTAGAATAATGGATGGTCTTGGATGATGGATAAA  
TAGGGACAGGGACAGTTAAATTGGGAGCCTTTCTTACAACCTTGATGGGATTTTTCCCCC  
CAAGTTTCCTTCTCCACTGAAATGCCACACTAATGCTTGTTGGATTCATGAGGTGGCCAG  
ACCAATGTGTTGTTTTGTTGTTGTTTTTTTTTTAAGCTTCCCTTGAGAGAATAAATGGTA  
ATGGAGAGAATCATTTAACAAGGTCTGGTTTCTCTTGCAACACAGTAGCTAAACTTGCC  
TGCTTTTATATGCATTTTTGTAGGGATCAGCTTGGTAGACAGTATTAGCGGAGAAACACC  
TTGATCTTGGTTTGCAAGCCCTTCTCCCATCAGTCCTAGATTAGGCCCTGTTCAGCCATG  
CAGGGGTGTTGGTTTATGCGTGCTGCAGCAGTGGGCATAATGAATATAATTTACCCAGTG  
GACAAAGGTGTGTACCAAGTGAATTTAAATAATTGGTGTGGATTGGCCAGTAGCTAAGAA  
GTGGGCTTTTAAAGAGTATTGAAGATTGAAAGGGTTTTTCTTTCTTTTTTAAAAAAGAAA  
AACAACTATTGATTGTAGATAATGAAAAGCTAGGGTTTGCCCTCTTCATGTCCTACTCTC  
CTTCCAAATAGTTATATCCAAACTGTTTTTCCCTCTCCCCTACCTTGTCCTCCCTATTA  
AAATAGAAACAGGGATTGATTAATGTCCCGCTCCTGAATACATGTAAAAATTTGTACAAAA  
ATATCTTCTATGAAAATGATTTGTAATCTGTAGACTTATTACCTGGGAGATGTCTTGATG  
TAAATCCCATCCTTTGGGTGTGGGTTTTTTGTTTTCTCCAAATAAATCTGATCTTTAA  
AGTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA  
AGTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA

354/392

**FIGURE 340**

MAASAKKKKNGKGTISLTDFLAEDGGTGGGSGTYVSKPVSWADETDDLEGDVSTTWHSNDD  
DVYRAPPIDRSILPTAPRAAREPNIDRSRLPKSPPYTAFLGNLPYDVTEESIKEFFRGLN  
ISAVRLPREPSNPERLKGFYAEFEDLDSLLSALSLSNEESLGNRRIRVDVADQAQDKORD  
DRSFGRDRNRDSDKTDTDWRARPATDSFDDYPPRRGDDSFQDKYRDRYDSDRYRDGYRDG  
YRDGPRRDMDRYGGDRYDDRGSRDYDRGYDSRIGSGRRAFSGGYRRDDDYRGGGDRYED  
RYDRRDRSWSSRDDYSRDDYRRDDRGPQRPKLNLKPRSTPEEDDSSASTSQSTRAASI  
FGGAKPVDTAAREEREVEERLQKEQEKLRQWNEPKLERPRERHPSWRSEETQERERSRT  
GSESSQTGTSTTSSRNARRRESEKSLNETLNKEEDCHSPTSKPPKPDQPLKVMPPPPK  
ENAWVKRSSNPPPARSQSSDTEQQSPTSGGGKVAPAPQSEEGPGRKDENVKVDGMNAPKQQT  
GNSSRGPGDGGNRDHWKESDRKDGKKDQDSRSAPEPKPEENPASKFSSASKYAALSVDG  
EDENEGEDYAE

**Signal Sequence:**

None

**Transmembrane domain:**

None

**N-glycosylation site:**

120-123, 448-451, 542-545

**Glycosaminoglycan attachment site:**

507-510

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

439-442, 486-489

**Tyrosine kinase phosphorylation site:**

225-233, 264-270

**N-myristoylation site:**

25-30, 26-31, 28-33, 118-123, 421-426, 428-433, 538-543

**Amidation site:**

276-279, 522-525, 563-566

**Cell attachment sequence:**

215-217

**Eukaryotic putative RNA-binding region RNP-1 signature:**

137-144

**RNA recognition motif:**

98-168

355/392

**FIGURE 341**

CGGTGGACACCACCTCAGCCCCTGAGCAGGAGTCACAGCACGAAGACCAAGCGCAAAGC  
GACCCCTGCCCTCCATCCTGACTGCTCCTCCTAAGAGAGATGGCACC GGCCAGAGCAGGA  
TTCTGCCCCCTTCTGCTGCTTCTGCTGCTGGGGCTGTGGGTGGCAGAGATCCCAGTCAGT  
GCCAAGCCCAAGGGCATGACCTCATCACAGTGGTTTAAAATTCAGCACATGCAGCCCAGC  
CCTCAAGCATGCAACTCAGCCATGAAAAACATTAACAAGCACACAAAACGGTGCAAAGAC  
CTCAACACCTTCCTGCACGAGCCTTTCTCCAGTGTGGCCGCCACCTGCCAGACCCCCAAA  
ATAGCCTGCAAGAATGGCGATAAAAACTGCCACCAGAGCCACGGGCCCCGTGTCCCTGACC  
ATGTGTAAGCTCACCTCAGGGAAGTATCCGAAGTGCAGGTACAAAGAGAAGCGACAGAAC  
AAGTCTTACGTAGTGGCCTGTAAGCCTCCCCAGAAAAAGGACTCTCAGCAATTCCACCTG  
GTTCTGTACACTTGGACAGAGTCCTTTAGGTTTCCAGACTGGCTTGCTCTTTGGCTGAC  
CTTCAATTCCCTCTCCAGGACTCCGCACCACTCCCCTACACCCAGAGCATTCTCTTCCCC  
TCATCTCTTGGGGCTGTTCTGGTTTCAGCCTCTGCTGGGAGGCTGAAGCTGACACTCTGG  
TGAGCTGAGCTCTAGAGGGATGGCTTTTCATCTTTTGTGCTGTTTTCCAGATGCTTA  
TCCCCAAGAAACAGCAAGCTCAGGTCTGTGGGTTCCCTGGTCTATGCCATTGCACATGTC  
TCCCCTGCCCCCTGGCATTAGGGCAGCATGACAAGGAGAGGAAATAAATGGAAAGGGGGC  
AAA  
AAA

356/392

## FIGURE 342

MAPARAGFCPLLLLLLLGLWVAEIPVSAKPKGMTSSQWFKIQHMQPSPQACNSAMKNINK  
HTKRCKDLNTFLHEPFSSVAATCQTPKIACKNGDKNCHQSHGFPVSLTMCKLTSGKYPNCR  
YKEKRQNKSYVVACKPPQKKDSQQFHLVPVHLDRVL

### Important features of the protein

**Signal peptide:**

1-22

**Transmembrane domain:**

none

**N-glycosylation site:**

127-131

**cAMP- and cGMP-dependent protein kinase phosphorylation  
site:**

139-143

**N-myristoylation site:**

18-24, 32-38

**Pancreatic ribonuclease family signature:**

65-72

**Pancreatic ribonuclease family proteins:**

49-93

357/392

**FIGURE 343**

GCATTTGCCACTGGTTGCAGATCAGGCGGACGAGGAGCCGGGAAGGCAGAGCCATGTGGC  
TGGCCCCCTGCTCTGCTCCTTCTCAGCCTCTCAGGCTGTTTCTCCATCCAAGGCCAGAGT  
CTGTGAGAGCCCCAGAGCAGGGGTCCCTGACGGTTCAATGCCACTATAAGCAAGGATGGG  
AGACCTACATTAAGTGGTGGTGCCGAGGGGTGCGCTGGGATACATGCAAGATCCTCATTG  
AAACCAGAGGGTCCGAGCAAGGAGAGAAGAGTGACCGTGTGTCCATCAAGGACAATCAGA  
AAGACCGCACGTTCACCTGTGACCATGGAGGGGCTCAGGCGAGATGACGCAGATGTTTACT  
GGTGTGGGATTGAAAGAAGAGGACCTGACCTTGGGACTCAAGTGAAAGTGATCGTTGACC  
CAGAGGGAGCGGCTTCCACAACAGCAAGCTCACCTACCAACAGCAATATGGCAGTGTTCA  
TCGGCTCCCACAAGAGGAACCACTACATGCTCCTGGTATTTGTGAAGGTGCCATCCTTGC  
TCATCTTGGTCACTGCCATCCTCTGGTTGAAGGGTCTCAGAGGGTCCCTGAGGAGCCAG  
GGGAACAGCCTATCTACATGAACTTCTCCGAACCTCTGACTAAAGACATGGCCACTTAGA  
GAGATGGATCTGCAGAGCCTTCTGCCCTGGCCACGTTTCCAGAAGAGACTCGGGCTGTG  
GAAGGAACATCTACGAGTCTCGGGATGCAGTGACTGAGATAGGGGCCCTGGGCCTCCGC  
CCTGGCCTTGGAGCTGGTGGGCACCTCCCTGTTCTGCACAGCTCAGGGACTTAGCCAGGT  
CCTCTCCTGAGCCACCATCACCTCCTGGGGTGCCAGCACCTGTTCTCTTGGTCAGGAGCT  
GTAGAGATGGAGCTCAAGCACTGGACGACTCTGTCCCCACTGCTGGAATAACTCGGGCAC  
AGAGCATGGGACCAAAGTACAGAAAGAGGTTGGGGGAGACCCCCCAGCCCTAGACTTCC  
ATCATTCGGGAGACCAACTCAACACCGTCTTTGCCCTGAGAACCTGATATATCCGTGTTTT  
TAAATTTTTTTTTTTCTAGCAAAGTTGGGTTTTAATGACTTATGTTTCATAGGAAACCTCT  
CTGATCCCACACACAAGGAGGGTGATTCTGGGATGAGTTCTGTTCTAGGGCATGAGGG  
GCTGGATGGACCCTGTCCCCAGGGAGGACATGGCTCTGAGTCCACAGGGCTGAGGAGGCA  
ATGGGAACCTCCCTGGCCCCGGCCGGTGCTTGTCTCCCCCTCCACCTCTTCTCCTCC  
TAGCTCCCCAAGCTCCCTGCCTATTCCCCACCTCCGAGGGGCTGCAGCTTGGGAGCCTC  
CTCAGCATGACAGCTTGGGTCTCCTCCCCAAAAGAGCCTGTCAGGCCTCAAGAACCACCT  
CCAGGTGGGGAGGGCAGTAACGAAAACCATCGCAGGAAATGGCACCCCTCCCTTTTCGGTG  
ATGTTGAAATCATGTTACTAATGAAAACCTGTCTAGGGAAGTGGTCTGTCTCCTCACAG  
GCTTCACCCACGGCGATGAGGCCCTTGAATGTGGTCACTTTGTGCTGTATGGTTGAGGGA  
CCCTCACACCAAAGGGACCTTCCCATGTGAGATGTGCTCCCGCCCCACCTGCCACAAG  
CAAACACACCACACATGTTCCGCATGTTGCCCTTTGAACACCCATGAGGACGCCTCCAAC  
CTGCTCTTGGTTCTAATAGGGAGTACTGACTGTGAGCAGTGGATAAAGGAGAGGGGACCC  
TCTGGTCCCTAGCATGGCACCCAGAGCCTCCCCTCTTCTTGTCTTCAGCCAAAGAGAAA  
CTTTCTCTGACTTTGAACTGAATTTAGGTCTCTGGCCAATGATGGGCCTGAAAATCCAT  
AATGGCCAGAGAGGAGAGTTCGAGCCCGGCTAAGATCCCCTGAGTCATTCTGTGAGGGAC  
CAAGACCCACAGTCCACCAGCCCCAGGGCCCTACCTCCTGGAATGCTTTCCTGGATCCAG  
CTTCCCGAAGATCCGACCAGACCCAGGGAGGACGGCACCGCTCCGCGGGAGGGAAAAGCCA  
AAGCATGGTGCTTACCAGCTGGAATCAGGGGCGAGGGGACATGGGCGCTTGTCAACGTG  
ATGTCATTCTTTTCCACCGTTTCTTCTGTTGATATTCAATGAATCCGTCAATCTCTCT  
GGGAAA

358/392

## FIGURE 344

MWLPPALLLLSLSGCFSSIQGPESVRAPEQGS�TVQCHYKQGWETIYIKWWCRGVRWDTCKI  
 LIETRGSEQGEKSDRVSIKDNQKDRTFTVTMEGLRRDDADVWCGIERRGPDLTQVKVI  
 VDPEGAASTTASSPTNSNMAVFIGSHKRNHYMLLVFVKVPILLILVTAILWLKGSQRVPE  
 EPGEQPIYMNFSPLTKDMAT

**Important features of the protein:**

**Signal peptide:**

Amino acids 1-17

**Transmembrane domain:**

Amino acids 151-170

**N-glycosylation site:**

Amino acids 190-194

**Tyrosine kinase phosphorylation site:**

Amino acids 95-103

**N-myristoylation sites:**

Amino acids 66-72;125-131

**Prokaryotic membrane lipoprotein lipid attachment site:**

Amino acids 5-16

359/392

## FIGURE 345

CTGAGCTCCCGGGCTCCGGCAGCGCGCTGGCGGGGCGCCGCATTGCACACTCTGGGGGCG  
CCGCAGTGTTCGTGGGATGGGGCAGCGGGCTGCAGCTGGCGGCCGGAATCCGCGCGCAGC  
CCGGGTGCAAGTTCTCTCCTGTTGCCCTGAGTGCCCACTCCCAGGCCCTCTGTATGAGTG  
ACACTTCAGTCTGCCATGGAACCTGGCCCTGCTCTGGCCTGGCTCCTGCTCCTGAGCCTG  
CTGGCGGATTGTCTGAAAGCTGCTCAGTCCCAGACTTCACAGTGAAAGACATTATCTAC  
CTCCATCCTTCAACCACACCATATCCTGGTGGATTTAAATGTTTCACCTGTGAAAAGGCA  
GCAGACAATTATGAGTGCAACCGATGGGCTCCAGACATCTACTGCCCTCGAGAGACCAGA  
TACTGCTACACTCAGCACACAATGGAAGTCACAGGAAACAGTATCTCAGTCACCAAACGC  
TGTGTCCCACTGGAAGAGTGCTTATCCACTGGCTGCAGAGACTCCGAGCATGAAGGCCAC  
AAGGTCTGCACCTTCTTGTGTGAAGGAATATCTGTAACCTGCCACTGCCCCGAAATGAA  
ACTGATGCCACATTTGCCACGACGTACCTATAAATCAGACAAATGGGCACCCACGCTGT  
ATGTCAGTGATAGTGTCTGCTTGTGGTTGTGGTTAGGGCTCATGTTATAGTGGCTCAGT  
GGCTCCATGTGTTAATAGCGATCCATGGGGATCTCGATGGTCCACAGACCTGCATGAGTC  
ATTGGCCTGACAGTAATTACACATGTGAGACACAACACTCTTGGAGGTCATCACAGCCAA  
GCATTGCCACTTACCATGAGGAATAAATGTTGCTTCATTGTAGCCATTTTGAGCTAACC  
GAGACTCATCAAAGCCTTCTGTCTAGTACAGCCCAAGTTCCATACCATAAACGTTGTTTT  
CATTCCAAGAAGTAGTTCTGCATTTATCGAGATCTGGGGTTCTTAATTTGGAAGAATACA  
TGCATGAGATGCAGTAGGTCTGAGACTGTAAGATATTAGGAGTATGTTATAGGGGCATG  
TATAGATGTGGGCTTTTCAGGAGAAAAGTAACCATTGGTTTAAATATAATCATGAGTTCA  
TTTGTAGCTTTAGAAATTTTAAACATTGACTCCAACTGAATGGACTATTTCTTGGA  
TTCTGACTGAGTCCCTGGAAGAGTAGTAATTCACAATTCAGCCATTTGTTCAATTAA  
TTTTCCCAACATTTCTTCTCCAGTGCTGGGAATCACATTTCTCTGTTCTGTGCAGAAGA  
CAAAAAGGCAATCATAAAAGTTTGTATATTTGTGGGGGTGCCTGGAGGAGGATTTTCT  
CAACTTAATGGAGCCACTGTCCATAAAGTGGCTGTTATCCCTTCATATAATTGGTGAGAT  
CAGCCTTCTCCTTGACTTGGCACCTAATTATGCTTCATGAGATCCTAGATTCCACCTGAG  
TCAATTGTGTCCAGAGCCCCAAACCAGGATGGAGTTGTTTTCCCAGATATGGGGTTCTA  
TTCAGCCATAGATAATCTAGACAGAGGATTTCAGAAATGAAAGGAAAAATGTGTGGAGATT  
AGTCCTAGTTTCTGAGGGCCGACTAAGTGGCTCAGCCAGCTTCTTACTCCATCTGCA  
GTTTCACTAGTCCAAAGAGCTCCCACTTCCAAATCCCAGTGACTTTATGGAGAAGATTCT  
GCATTAAATTGTCTTTTGAATGATGGGGAAGCAAGGCATAATATGCGATGATGAGGAGAA  
AGTAGACCAGTGAGGTGATTGCAAGACTAACAAGGAGACTCAATGGGAAGTTTTTCTTCT  
TTTATAGATATTGCTTTTGAAGTAGATGGTAAAATTTTGTATCCTTCTGTATTTTTTG  
TACCCCAAGTTACAATTTTTCTTCTTCTTGTAAATAATTTAAACAGTATTTATTTTTGT  
AAGGCATAACTAGAACTAAAATATATTCTAAAAAATTCATTATTCTGAACAAAGTGATC  
AAATTAGAATACATATTTTTCAACAGTGGTAGAGCTTTTAATATATGTTTATTGAAAGTT  
ATCTATAATACTTGCACCAGTGTTGAAAAAAGTTAACATGTAGGCAAGAGCAATATGTTT  
GTCTCAAGGATTTTTCCATGGTTTTCTCAGTGATGGTGTCTGGAATTATTAGGTGGTG  
ACCATCACTGGTCTAAGTTTGTGTGCAGGGTTTTTCAGACGTGTTTTTGTGAACTTGGTA  
GAACCATGGCTAATAAAGAGGACAGTGTTGTGAGGGTCCATCTGCCCTCCATAGAAAAAT  
GTCTCTGGCTCATAAAATGAGACTCCCTCAGGGACTAAATATGAACTGACAGCAGTAACT  
CTGATACAGAATAATCTAAATTGCATCAAATGGCCTTAATTCAGAGTTTGTAGGCTTAT  
CAGTATGTTGCTTTTAAATTGGGGTGGGAAAGTAGAGGGAGAGAAAGCAAGACATTTATTA  
AGCACCTCGTATGTGCCAGGCACTATGCTAAGCACTTTACATAAGTTAGGATTAATCCCT  
GCAAGAATCCTATAAAGAATGTTACTAGCATTTTACACTTCCCAATGAAGGTACCAAAGC  
TCAAACGCAATGTTGTGAAGCTGTTTCTTCTCAGATTTAGGTTATGTGGGATGATGTGGGA  
TTGAAGAGGAAAGAAAGGTGGGATTATCCCCCTAGGAAGACTTTTCCAGCCTGACTTCATA  
GGAATTCATCCATCTTATCATGTGGAGTTTATCTCACCCTGCTGTTGCAGGATGCTATTT  
GCATGTGTCCCAGGTGATGTTTTTTCTTGGGGAGTAGGGTTTTGGCTTCTCATTTCAT

360/392

CCCTCTGCTAAAAGAGGAGATAGTTGATGTTGCATCTAAGATGCTATAAGCAATGAAAGTTTGATGTTGTAC  
ATACCTACAAGTACCATTTTTTGTGCATGATTACACTCCACTGACATCTTCCAAGTACTAC  
ATGTGATTGAATAAGAAACAAGAAAGTGACCACACCAAAGCCTCCCTGGCTGGTGTACAG  
GGATCAGGTCCACAGTGGTGCAGATTCAACCACCACCCAGGGAGTGCTTGCAGACTCTGC  
ATAGATGTTGCTGCATGCGTCCCATGTGCTGTGAGAAATGGCAGTGTTTAATTCTCTTGA  
AAGAAAGTTATTTGCTCACTATCCCCAGCCTCAAGGAGCCAAGGAAGAGTCATTCACATG  
GAAGGTCCGGGACTGGTCAGCCACTCTGACTTTTCTACCACATTAAATTCTCCATTACAT  
CTCACTATTGGTAATGGCTTAAGTGTAAGAGCCATGATGTGTATATTAAGCTATGTGCC  
ACATATTTATTTTTTAGACTCTCCACAGCATTCATGTCAATATGGGATTAATGCCTAAACT  
TTGTAAATATTGTACAGTTTGTAATCAATGAATAAAGGTTTTGAGTGTAACAAAAAAAAAA  
AAAAAA



361/392

## FIGURE 346

MEPGPALAWLLLLSLADCLKAAQSRDFTVKDIIYLHPSTTPYPGGFKCFTCEKAADNYE  
CNRWAPDIYCPRETRYCYTQHTMEVTGNSISVTKRCVPLEECLSTGCRDSEHEGHKVCTS  
CCEGNICNLPLPRNETDATFATTSPINQTNHGPRCMSVIVSCLWLWLGLML

Important features of the protein:

Signal peptide:

1-22

Transmembrane domain:

None

N-glycosylation site:

134-138, 147-151

N-myristoylation site:

45-51, 87-93, 106-112, 124-130

Ly-6 / u-PAR domain protein:

115-128

362/392

**FIGURE 347**

GATCAAGCGCCTTCCTTTCCCTTCCTCTCCCTACTTGGCCTTTGCCCTAAGCCAAGACCT  
GGCCATCAGCCTGGCTGCAGGGGCTGCAGAGCCAGCTGCACTTTTTCAGGTATGGGGGA  
GGGCCAGGCACCATGAAGCCAGTGTGGGTGCCACCCTTCTGTGGATGCTACTGCTGGTG  
CCCAGGCTGGGGGCCGCCCGGAAGGGGTCCCAGAAGAGGCCTCCTTCTACTATGGAACC  
TTCCCTCTTGGCTTCTCCTGGGGCGTGGGCAGTTCCTGCCTACCAGACGGAGGGCGCCTGG  
GACCAGGACGGGAAAGGGCCTAGCATCTGGGACGTCTTACACACAGTGGGAAGGGGAAA  
GTGCTTGGGAATGAGACGGCAGATGTAGCCTGTGACGGCTACTACAAGGTCCAGGAGGAC  
ATCATTCTGCTGAGGGAAGTGCACGTCAACCACTACCGATTCTCCCTGTCTTGGCCCCGG  
CTCCTGCCCACAGGCATCCGAGCCGAGCAGGTGAACAAGAAGGGAATCGAATTCTACAGT  
GATCTTATCGATGCCCTTCTGAGCAGCAACATCACTCCCATCGTGACCTTGCACTACTGG  
GATCTGCCACAGCTGCTCCAGGTCAAATACGGTGGGTGGCAGAATGTGAGCATGGCCAAC  
TACTTCAGAGACTACGCCAACCTGTGCTTTGAGGCCCTTTGGGGACCGTGTGAAGCACTGG  
ATCACGTTCAGTGATCCTCGGGCAATGGCAGAAAAAGGCTATGAGACGGGGCCACCATGCG  
CCGGGCCTGAAGCTCCGCGGCACCGGCCTGTACAAGGCAGCACACCACATCATTAAGGCC  
CACGCCAAAACCTGGCATTCTTATAACACCACGTGGCGCAGCAAGCAGCAAGGTCTGGTG  
GGAATTTCACTGAACTGTGACTGGGGGGAACCTGTGGACATTAGTAACCCCAAGGACCTA  
GAGGCTGCCGAGAGATACCTACAGTTCTGTCTGGGCTGGTTTGCCAACCCCATTTATGCC  
GGTGACTACCCCAAGTCATGAAGGACTACATTGGAAGAAAGAGTGCAGAGCAAGGCCTG  
GAGATGTCGAGGTTACCGGTGTTCTCACTCCAGGAGAAGAGCTACATTAAAGGCACATCC  
GATTTCTTGGGATTAGGTCAATTTACTACTCGGTACATCACGAAAGGAACTACCCCTCC  
CGCCAGGGGCCAGCTACCAGAACGATCGTGACTTGATAGAGCTGGTTGACCCAAACTGG  
CCAGATCTGGGGTCTAAATGGCTATATTCTGTGCCATGGGGATTAGGAGGCTCCTTAAC  
TTTGCTCAGACTCAATACGGTGATCCTCCCATATATGTGATGGAAAATGGAGCATCTCAA  
AAATTCCACTGTACTCAATTATGTGATGAGTGGAGAATTCAATACCTTAAAGGATACATA  
AATGAAATGCTAAAAGCTATAAAAGATGGTGCTAATATAAAGGGGTATACTTCCTGGTCT  
CTGTTGGATAAGTTTGAATGGGAGAAAGGATACTCAGATAGATATGGATTCTACTATGTT  
GAATTTAACGACAGAAATAAGCCTCGCTATCCAAAGGCTTCAGTTCAATATTACAAGAAG  
ATTATCATTGCCAATGGGTTTCCCAATCCAAGAGAGGTGGAAAGTTGGTACCTCAAAGCT  
TTGGAAACTTGCTCTATCAACAATCAGATGCTTGCTGCAGAGCCTTGCTAAGTCACATG  
CAAATGGTTACGGAGATCGTGGTACCCACTGTCTGCTCCCTCTGTGTCCTCATCACTGCT  
GTTCTACTAATGCTCCTCCTGAGGAGGCAGAGCTGAGACAGGATTATCAATTTTGAGCT  
TCATAAGAGAATCTTCAGGATCTTCCTCCCTTTTCTGCTTTGAGGGTTTCCATACATTGC  
TGTTTTTCAGGTTCTACAATAATTACCTTTTTTTCTCTTTCTCTTTTGGCTTGTGCTGGG  
ATTTAAGAATTAGAAAATAAAAATAAGCAGAAATTA

363/392

**FIGURE 348**

MKPVVWVATLLWMLLLVPRLGAARKGSPEEASFYYGTFPLGFSWGVGSSAYQTEGAWDQDG  
KGPSIWDVFTHSGKGKVLGNETADVACDGYKQVEDIILLRELHVNHYRFSLSWPRLPT  
GIRAEQVNKKGIEFYSDLIDALLSSNITPIVTLHHWDLPQLLQVKYGGWQNVSMANYFRD  
YANLCFEAFGDRVKHWITFSDPRMAEKGYETGHHAPGLKLRGTGLYKAAHHIIKAHAKT  
WHSYNTTWRSKQQGLVGLISLNCDWGEPVDISNPKDLEAAERYLQFCLGW FANPIYAGDYP  
QVMKDYIGRKSAEQGLEMSRLPVFSLQEKSYIKGTSDFLGLGHFTTRYITERNYPSRQGP  
SYQNDRLIELVDPNWPDLGSKWLYSVPWGFRLLNFAQTQYGDPPYVMENGASQKFHC  
TQLCDEWRIQYLKGYINEMLKAIKDGANIKGYTSWSLLDKFEWEKGYSDRYGFYYVEFND  
RNKPRYPKASVQYYKKIIANGFPNPREVESWYLKALETCSINNQMLAAEPLL SHMQMVT  
EIVVPTVCSLCVLITAVLLMLLLRRQS

**Important features:****Signal peptide:**

amino acids 1-21

**Transmembrane domain:**

amino acids 541-558

**N-glycosylation sites:**

amino acids 80-84,171-175,245-249

**Glycosaminoglycan attachment site:**

amino acids 72-76

**cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

amino acids 23-27,564-568

**Tyrosine kinase phosphorylation sites:**

amino acids 203-211,347-355,460-468,507-514

**N-myristoylation sites:**

amino acids 44-50,79-85,167-173,225-231,257-263,315-321

**Amidation site:**

amino acids 307-311

**Glycosyl hydrolases family 1 active site:**

amino acids 407-416

**Glycosyl hydrolases family 1 N-terminal signature:**

amino acids 41-56

**Motif name Glycosyl hydrolases family:**

amino acids 37- 67

364/392

**FIGURE 349**

CGCAAAGCCGCCCTCGGGGCGCTCATGGCGGGACGCCTCCTGGGAAAGGCTTTAGCCGCG  
GTGTCTCTCTCTCTGGCCTTGGCCTCTGTGACTATCAGGTCCTCGCGCTGCCGCGGCATC  
CAGGCGTTTCAAGAACTCGTTTTTCATCTTCTTGGTTTCATCTTAATACCAACGTCATGTCT  
GGTTCTAATGGTTCCAAAGAAAATTCTCACAATAAGGCTCGGACGTCTCCTTACCCAGGT  
TCAAAGTTGAACGAAGCCAGGTTCCCTAATGAGAAAGTGGGCTGGCTTGTTGAGTGGCAA  
GACTATAAGCCTGTGGAATACACTGCAGTCTCTGTCTTGGCTGGACCCAGGTGGGCAGAT  
CCTCAGATCAGTGAAAGTAATTTTTCTCCCAAGTTTAACGAAAAGGATGGGCATGTTGAG  
AGAAAGAGCAAGAATGGCCTGTATGAGATTGAAAATGGAAGACCGAGAAATCCTGCAGGA  
CGGACTGGACTGGTGGGCCGGGGCTTTTTGGGGCGATGGGGCCCAAATCACGCTGCAGAT  
CCCATTTATAACCAGATGGAAAAGGGATAGCAGTGGAAATAAAATCATGCATCCTGTTTCT  
GGGAAGCATATCTTACAATTTGTTGCAATAAAAAGGAAAAGACTGTGGAGAATGGGCAATC  
CCAGGGGGGATGGTGGATCCAGGAGAGAAGATTAGTGCCCACTGAAAAGAGAATTTGGT  
GAGGAAGCTCTCAACTCCTTACAGAAAACCAGTGCTGAGAAGAGAGAAAATAGAGGAAAAG  
TTGCACAAACTCTTCAGCCAAGACCACCTAGTGATATATAAGGGATATGTTGATGATCCT  
CGAAACACTGATAATGCATGGATGGAGACAGAAGCTGTGAACTACCATGACGAAACAGGT  
GAGATAATGGATAATCTTATGCTAGAAGCTGGAGATGATGCTGGAAAAGTGAAATGGGTG  
GACATCAATGATAAACTGAAGCTTTATGCCAGTCACTCTCAATTCATCAAACTTGTGGCT  
GAGAAACGAGATGCACACTGGAGCGAGGACTCTGAAGCTGACTGCCATGCGTTGTAGCTG  
ATGGTCTCCGTGTAAGCCAAAGGCCACAGAGGAGCATATACTGAAAAGAAGGCAGTATC  
ACAGAATTTTATACTATAAAAAGGGCAGGGTAGGCCACTTGGCCTATTTACTTTCAAAACA  
ATTTGCATTTAGAGTGTTTCGCATCAGAATAACATGAGTAAGATGAACTGGAACACAAAA  
TTTTTCAGCTCTTTGGTCAAAAGGAATATAAGTAATCATATTTTGTATGTATTCGATTTAA  
GCATGGCTTAAATTTAAATTTAAACAATAATGCTCTTTGAAGAATCATAATCAGAATAAA  
GATAAATCTTGATCAGCTATA

365/392

**FIGURE 350**

MAGRLLGKALAAVSLSLALASVTIRSSRCRGIQAFRNSFSSSWFHLNTNVMMSGNNGSKEN  
SHNKARTSPYPGSKVERSQVPNEKVGWLVEWQDYKPVEYTAVSVLAGPRWADPQISESNF  
SPKFNEKDGHVERKSKNGLYEIENGRPRNPAGRTGLVGRGLLGRWGPNAADPIITRWKR  
DSSGNKIMHPVSGKHILQFVAIKRKDCGEWAI PGGMVDPGEKISATLKREFGEEALNSLQ  
KTSAEKREIEEKLHKLFSQDHLVIYKGYVDDPRNTDNAWMETEAVNYHDETGEIMDNMLL  
EAGDDAGKVKWVDINDKLKLYASHSQFIKLVAEKRDAHWESEADCHAL

**Important features of the protein:****Signal peptide:**

1-20

**Transmembrane domain:**

None

**N-glycosylation site:**

55-59

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

179-183

**N-myristoylation site:**

53-59, 56-62

**mutT domain signature:**

215-235

366/392

**FIGURE 351**

CCTCTGTCTGTGCTCCCATCCCAGGGAGTATAGGTGGAGCCTCCAGAGCCCATGGACAGG  
GCATGCTGGGGCTGGGCCAGCCCCAGCGGTGTCTCTAAGGCACCCCTGGGATCCCCACTG  
AGCTGGCCTACTTCAGACAGCCAGGGCCCCACCCCTCTGGCCCCCTTAGTGTCAGCTCGT  
GGCCCCCTTGGCATTTCCACAAGACGCCAAGATGGAGATTCCCATGGGGACCCAGGGCTGC  
TTCTCAAAGAGCCTCCTGCTCTCAGCCTCAATCCTGGTCCTCTGGATGCTCCAAGGCTCC  
CAGGCAGCTCTCTACATCCAGAAGATTCCAGAGCAGCCTCAAAAGAACCAGGACCTTCTC  
CTGTTCAGTCCAGGGTGTCCCAGACACCTTCCAGGACTTCAACTGGTACCTGGGGGAGGAG  
ACGTACGGAGGCACGAGGCTATTTACCTACATCCCTGGGATACAACGGCCTCAGAGGGAT  
GGCAGTGCCATGGGACAGCGAGACATCGTGGGCTTCCCCAATGGTTCCATGCTGCTGCGC  
CGCGCCAGCCTACAGACAGTGGCACCTACCAAGTAGCCATTACCATCAACTCTGAATGG  
ACTATGAAGGCCAAGACTGAGGTCCAGGTAGCTGAAAAGAATAAGGAGCTGCCAGTACA  
CACCTGCCCCACCAACGCTGGGATCCTGGCGGCCACCATCATTTGGATCTCTTGCTGCCGGG  
GCCCTTCTCATCAGCTGCATTGCCTATCTCCTGGTGACAAGGAAGTGGAGGGCCAGAGC  
CACAGACTGCCTGCTCCGAGGGGCCAGGGATCTCTGTCCATCTTGTGCTCGGCTGTATCC  
CCAGTGCCTTCAGTGACGCCCAGCACATGGATGGCGACCACAGAGAAGCCAGAAATTGGGC  
CCTGCTCATGATGCTGGTGACAACAACATCTATGAAGTGATGCCCTCTCCAGTCCTCCTG  
GTGTCCCCCATCAGTGACACAAGGTCCATAAACCCAGCCCGGCCCTGCCACACCCCCA  
CACCTGCAGGCGGAGCCAGAGAACCACAGTACCAGCAGGACCTGCTAAACCCCGACCCT  
GCCCCCTACTGCCAGCTGGTGCCAACTTCCTGATGGGTCTTGGGCCAGGCCAGCCAGGGA  
GAAGACAAGGCCCCAGCCCTCCTCTGGGAGCCTCACACCTGAGACCAGCAGGACAAGGCC  
ATTGGGGGCTGTGGGGCCGATGAGGTGGACTCAGCCAAAGACTCAGCAGCACATGGGGCA  
GGTGTCTTGGCAGGGGGACAGGAGACTGTAACAGGCCCAGGTCTTGTGCAGCCCCGTGAA  
TGACGCCCCGCTTCGGTCTGTTCTTCAAGCAAGCTGGCCTGGGCCATGTGCCCTGTGAA  
AGGCAGGCTCTGGCCCCCTTCCATGCCAAAGTCCCCCAAGATCTGGATATCTGGGGACAA  
GATGGTGGCCTCAGGCCTGCCTCCCAGGCAGTTGGCTGGGCTCCCAACTGTCTGTCTCTCA  
ATGCCCTACCCCAACTCCACTAGTGACCCTCAGAGTCTTCTCCCCCTTAGGACAAGGCAGA  
CACCCCACCATGCGGGCCTCAGGTGGCAGAGAGGCCAGCCTCACAGGCCTGTGGCCCCA  
CACACCAGTCCCAGCAAGGTGACCACGGCTGCTGGACCCCTTCCCTGTTTCAGGCAGGCCC  
AGCCCCCTCTCAGAACCTGCTGCCAGCTGCTGGTCTTGGCCCCCACCCTGAATCTTACTGA  
GTCCCTCTGGGCAGCAGCTCCCTTCTCCACCCACCCAGCACCCGTCCCAAATGTGGCC  
TCAGCTTGTCTCTCCCTTCCCCAAACTATGCATTCAATCAGCAATAAATGAGCCTTTGCT  
GCA

367/392

## FIGURE 352

MEIPMGTOGCFKSLLLSASILVLWMLQGSQAALYIQKIPEQPQKNQDLLLSVQGVPDF  
QDFNWYLGEETYGGTRLFTYIPGIQRPQRDGSAMGQRDIVGFNGSMLLRRQAQPTDSGT  
QVAITINSEWTMKAKTEVQVAEKNKELPSTHLPTNAGILAATIIGSLAAGALLISCIAYL  
LVTRNWRGQSHRLPAPRGQGSLSILCSAVSPVPSVTPSTWMATTEKPELGPAGHDAGDNNI  
YEVMPSPVLLVSPISDTRSINPARPLPTPPHLQAEPENHQYQQDLLNPDPAPYCQLVPTS

**Important features of the protein:**

**Signal peptide:**

Amino acids 1-32

**Transmembrane domain:**

Amino acids 159-178

**N-glycosylation site:**

Amino acids 104-108

**N-myristoylation sites:**

Amino acids 6-12; 29-35; 55-61; 91-97; 157-163; 165-171

368/392

**FIGURE 353**

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAAGTTGCTCGAGTTAGAATTGTCTGCAATG  
GCCGCCCTGCAGAAATCTGTGAGCTCTTTCCTTATGGGGACCCTGGCCACCAGCTGCCTC  
CTTCTCTTGGCCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCACTGCAGG  
CTTGACAAGTCCAAC'TCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAG  
GAGGCTAGCTTGGCTGATAACAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCAC  
GGAGTCAGTATGAGTGAGCGCTGCTATCTGATGAAGCAGGTGCTGAAC'TTCACCC'TTGAA  
GAAGTGCTGTTCCCTCAATCTGATAGGTTCCAGCCTTATATGCAGGAGGTGGTGCCCTTC  
CTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCATATTGAAGGTGATGACCTGCATATC  
CAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGCTTGGAGAGAGTGGAGAGATC  
AAAGCAATTGGAGAACTGGATTGCTGTTTATGTCTCTGAGAAATGCCTGCATTTGACCA  
GAGCAAAGCTGAAAAATGAATAACTAACCCCTTCCCTGCTAGAAATAACAATTAGATG  
CCCCAAAGCGATTTTTTTTAACCAAAGGAAGATGGGAAGCCAACTCCATCATGATGGG  
TGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTTGTTTATA  
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTTCAATTGTAAC'TGG  
TGTTCTATACACAGAAAACAATTTATTTTTTAAATAATTGTCTTTTCCATAAAAAAGAT  
TACTTTCATTCCTTTtaggggAAAAAACCCTAAATAGCTTCATGTTTCCATAATCAGTA  
CTTTATATTTATAAATGTATTTATTATTATTATAAGACTGCATTTTATTTATATCATTTT  
ATTAATATGGATTTATTTATAGAAACATCATTGATATTGCTACTTGAGTGTAAGGCTAA  
TATTGATATTTATGACAATAATTATAGAGCTATAACATGTTTATTTGACCTCAATAACA  
CTTGGATATCCC



369/392

## FIGURE 354

MAALQKSVSSFLMGTLATSCLLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLA  
KEASLADNNTDVRLIGEKLFGVSMSERCYLMKQVLNFTLEEVLPQSDRFQPYMQEVVP  
FLARLSNRLSTCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

**Important features of the protein:**

**Signal peptide:**

amino acids 1-33

**N-glycosylation sites:**

amino acids 54-58, 68-72, 97-101

**N-myristoylation sites:**

amino acids 14-20, 82-88

**Prokaryotic membrane lipoprotein lipid attachment site:**

amino acids 10-21

370/392

**FIGURE 355**

TGGCCTACTGGAAAAAAAAAAAAAAAAAAGTCACCCGGGCCCGCGGTGGCCACAA  
CATGGCTGCGGCGCCGGGGCTGCTCTTCTGGCTGTTTCGTGCTGGGGGCGCTCTGGTGGGT  
CCCGGGCCAGTCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGGA  
CGAAGAGTGCAGCATGTTAATGTACCGTGGAAGCTCTTGAAGACTTCACGGGCCCTGA  
TTGTCGTTTTGTGAATTTTAAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGG  
GGGATCCCTTGAACTTTGGGCTGGAAGTGTTGAACACAGTTTGGATATTTTCCAAAAGA  
TTTGATCAAGGTACTTCATAAATACACGGAAGAAGAGCTACATATTCAGCAGATGAGAC  
AGACTTTGTCTGCTTTGAAGGAGGAAGAGATGATTTTAATAGTTATAATGTAGAAGAGCT  
TTTAGGATCTTTGGAAGTGGAGGACTCTGTACCTGAAGAGTCGAAGAAAGCTGAAGAAGT  
TTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGAACCTGACCCTGTGCC  
TGAGCCCGAGGCATTTCAGAGCTGATTTCAGAGGATGGAGAAGGTGCTTTCTCAGAGAGCAC  
CGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCTCACACCAGCGGTCCTGC  
GGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACACTTTTGAAGAAATTCTGCACGATAA  
ATTGAAAGTGCCGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCG  
GGAGAAGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACA  
GTGCGTTATTTCATTACAGCAAAGGATTTTCGTTGGCATCAAATCTAAGTTTGTTTTACAA  
AGATTGTTTTTAGTACTAAGCTGCCTTGGCAGTTTGCATTTTTTGAGCCAAACAAAAATAT  
ATTATTTTCCCTTCTAAGTAAAAAAAAAAAAAAAAAAAAA

371/392

## FIGURE 356

MAAAPGLLFWLFLVLGALWWVPGQSDLSHGRRFSDLKVCGDEEC SMLMYRGKALEDFTGPD  
 CRFVNFKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLH KYTEELHIPADET  
 DFVCFEGGRDDFNSYNVEELLGSLELEDSVPEESKKAEV SQHREKSPEESRGRELD PVP  
 EPEAFRADSEEDGE GAFSESTEG LQGQPSAQESH PHTSGPAANAQGVQSSLDTFEEILHDK  
 LKVPGESERTGNSSPASVEREKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNL SLYK  
 DCF

**Important features of the protein:**

**Signal peptide:**

amino acids 1-22

**N-glycosylation site:**

amino acids 294-298

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

amino acids 30-34

**Tyrosine kinase phosphorylation site:**

amino acids 67-76

**N-myristoylation sites:**

amino acids 205-211, 225-231, 277-283

**Amidation site:**

amino acids 28-32



373/392

## FIGURE 358

MEAPGPRALRTALCGGCCCLLLCAQLAVAGKGARGFGRGALIRLNIWPAVQGACKQLEVC  
EHCVEGDRARNLSSCMWEQCRPEEPGHCVAQSEVVKEGCSIYNRSEACPAHHHPTYPEK  
TVTTGSPVPPEAHSPGFDGASFIGGVVLVLSLQAVAFFVLHFLKAKDSTYQTL

### Important features of the protein:

#### Signal peptide:

1-29

#### Transmembrane domain:

141-160

#### N-glycosylation site:

71-75, 103-107

#### Tyrosine kinase phosphorylation site:

164-171

#### N-myristoylation site:

15-21

374/392

**FIGURE 359**

TTCCAGTCAGAGTTAAGTTAAAACAGAAAAAAGGAAGATGGCAAGAATATTGTTACTTTT  
CCTCCCGGGTCTTGTGGCTGTATGTGCTGTGCATGGAATATTTATGGACCGTCTAGCTTC  
CAAGAAGCTCTGTGCAGATGATGAGTGTGTCTATACTATTTCTCTGGCTAGTGCTCAAGA  
AGATTATAATGCCCCGGACTGTAGATTCATTAACGTTAAAAAAGGGCAGCAGATCTATGT  
GTA~~CT~~C~~CA~~AAGCTGGTAAAAGAAAATGGAGCTGGAGAATTTTGGGCTGGCAGTGT~~TT~~ATGG  
TGATGGCCAGGACGAGATGGGAGTCGTGGGT~~TATTT~~CCCCAGGA~~ACT~~TGGTCAAGGAACA  
GCGTGTGTACCAGGAAGCTACCAAGGAAGTTCC~~C~~ACCACGGATATTGACTTCTTCTGCGA  
GTA~~A~~A~~A~~A~~A~~ATTAGTTAA~~AA~~ACTGCAAATAGAAAGAAAACACCA~~AAAA~~ATAAAGAAAAGAGCAA  
AAGTGGCCAAAAAATGCATGTCTGTAA~~TTTT~~TGGACTGACGT

375/392

## FIGURE 360

MARILLFLPGLVAVCAVHGIFMDRLASKKLCADDECVTYISLASAQEDYNAPDCRFINV  
KKGQQIYVYSKLVKENGAGEFWAGSVYGDGQDEMGVVGYPFRNLVKEQRVYQEATKEVPT  
TDIDFFCE

Important features of the protein:

Signal peptide:

1-14

Transmembrane domain:

None

N-myristoylation site:

84-90

376/392

**FIGURE 361**

GGCACGAGCCACCACTTACAACCACACAGCCTATCCAGAAACATGAAGATAAGAAATGCT  
TGTGCTGTCTTATTGAAGTACTCCTGTTTATACTTGAAGGAGTTACAGGAGCTCGAAAA  
ATTTCAACTTTCTCAGGCCCTGGCTCATGGCCGTGCAATCCCAAGTGTGATGGCAGAACT  
TACAACCCCTCAGAGGAGTGTTGTGTTTCATGACACCATCCTGCCCTTTAAGCGGATTAAC  
CTCTGTGGCCCTAGCTGCACCTACAGGCCCTGCTTTGAGCTCTGCTGTCTGAGTCCTAT  
AGCCCCAAGAAGAAATTTATTGTCAAGCTTAAAGTTCATGGAGAGAGATCCCATTGCAGT  
TCATCCCCTATCTCCAGGAAGTGTAAAAGCAACAAGATTTTTTCATGGAGAAGATATTGAA  
GACAACCAACTTTCTCTTAGGAAAAAAGTGGTGACCAGCCTTGAGAGTCTGCTTTCTTC  
CTGCAAGCACCAGTTCCTGAATGTTCTTACTTGAAGAATGGATACCTGAAGCATTGGGGT  
GCAGTGATATATGTGTCTCATTACAATGCTCCTTTGGATATTGTTTTCTAAGCATGTGT  
TGGAATGTTCCCCATAACTTTCTAAAATTATCCTATTTCAATGCAACTAAAGATAAATG  
TATTCCAGCCAGAGTCCACAGAGAAGGCAAGTTATGCAAGGCAGGCATGGGGCCCTCACA  
AAATTTCAAGCTGTGCGACTTATGTAGTAATTTTCTACAAACAATCCCTCCTGGATATCC  
AGGAGGCTCCAGACCTGAATAAAAAACCACATGTCTGTCTAGAAAAAGGGAATGAATCAAG  
ATCCACAGGACCTTTTCAAGATTTTAGAAGCAGCAAACCTATGGCTGAGAGAAAAGACTCT  
CTGACCAGGCAAATTGTTCTGCAGTATTCTCCGGGCGTGTAGCTCCCCTGAGTAGTCGCC  
AGGCTGGTCTTGGCTTTGTAATAATACAGCTGCCTTTGAGTCCTCCCTACCCTGTTAGTA  
ACCCCTTGCTGCCTGCACTGTTGTCCTTACAACCGAAATAAACTGATTAGTTG



377/392

## FIGURE 362

MKIRNACAVLIEVLLFILEGVTGARKISTFSGPGSWPCNPKCDGRTYNPSEECVHDTIL  
PFKRINLCGPSCTYRPCFELCCPESYSPKKKFIVKLKVHGERSHCSSSPISRNCKSNKIF  
HGEDIEDNQLSLRKKSGDQP

### Important features of the protein:

#### Signal peptide:

1-23

#### Transmembrane domain:

None

#### Glycosaminoglycan attachment site:

31-35

#### N-myristoylation site:

20-26, 34-40

378/392

**FIGURE 363**

ACACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTC  
CCAGGACAGAGAGTGCACAACTACCCAGCACAGCCCCCTCCGCCCCCTCTGGAGGGCTGA  
AGAGGGATTCCAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCCCCTT  
GGGGGGGGGAGCACAGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAGATGCTCTGT  
GCCCTGGTTCTTGCTGTCTTGGCACTGGGCGGAAGCCCAGTGGTCCTTTCTCTGGAGAG  
GCTTGTGGGGCCTCAGGACGCTACCCACTGCTCTCCGGGCCTCTCCTGCCGCCTCTGGGA  
CAGTGACATACTCTGCCTGCCTGGGGACATCGTGCCCTGCTCCGGGCCCCGTGCTGGCGCC  
TACGCACCTGCAGACAGAGCTGGTGCTGAGGTGCCAGAAGGAGACCGACTGTGACCTCTG  
TCTGCGTGTGGCTGTCCACTTGGCCGTGCATGGGCACTGGGAAGAGCCTGAAGATGAGGA  
AAAGTTTGGAGGAGCAGCTGACTCAGGGGTGGAGGAGCCTAGGAATGCCTCTCTCCAGGC  
CCAAGTCGTGCTCTCCTTCCAGGCCTACCCTACTGCCCGCTGCGTCCTGCTGGAGGTGCA  
AGTGCCTGCTGCCCTTGTGCAGTTTGGTCAGTCTGTGGGCTCTGTGGTATATGACTGCTT  
CGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGTCCTATACTCAGCCCAGGTACGAGAA  
GGAACCTCAACCACACACAGCAGCTGCCTGCCCTGGCTCAACGTGTGACGAGATGG  
TGACAACGTGCATCTGGTTCTGAATGTCTCTGAGGAGCAGCACTTCGGCCTCTCCCTGTA  
CTGGAATCAGGTCCAGGGCCCCCAAAACCCCGGTGGCACAAAAACCTGACTGGACCGCA  
GATCATTACCTTGAACCACACAGACCTGGTTCCCTGCCTCTGTATTACAGGTGTGGCCTCT  
GGAACCTGACTCCGTTAGGACGAACATCTGCCCTTCAGGGAGGACCCCCGCGCACACCA  
GAACCTCTGGCAAGCCGCCGACTGCGACTGCTGACCCTGCAGAGCTGGCTGCTGGACGC  
ACCGTGCTCGCTGCCCGCAGAAGCGGCACTGTGCTGGCGGGCTCCGGGTGGGGACCCCTG  
CCAGCCACTGGTCCCACCGCTTTCCTGGGAGAACGTCACTGTGGACAAGGTTCTCGAGTT  
CCCATTGCTGAAAGGCCACCCTAACCTCTGTGTTACAGGTGAACAGCTCGGAGAAGCTGCA  
GCTGCAGGAGTGCTTGTGGGCTGACTCCCTGGGGCCTCTCAAAGACGATGTGCTACTGTT  
GGAGACACGAGGCCCCCAGGACAACAGATCCCTCTGTGCCTTGGAACCCAGTGGCTGTAC  
TTCACTACCCAGCAAAGCCTCCACGAGGGCAGCTCGCCTTGAGAGTACTTACTACAAGA  
CCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCCTG  
CCCCATGGACAAATACATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCCTGCCTACTCTT  
TGCCGCTGCGCTTTCCTCATCCTCCTTCTCAAAAAGGATCACGCGAAAGGGTGGCTGAG  
GCTCTTGAAACAGGACGTCCGCTCGGGGGCGGCCCGCAGGGGCCGCGCGGCTCTGCTCCT  
CTACTCAGCCGATGACTCGGGTTTCGAGCGCTGGTGGGCGCCCTGGCGTCCGCCCTGTG  
CCAGCTGCCGCTGCGCGTGGCCGTAGACCTGTGGAGCCGTGCTGAACCTGAGCGCGCAGGG  
GCCCCGTGGCTTGGTTTCACGCGCAGCGGCGCCAGACCCTGCAGGAGGGCGGCGTGGTGGT  
CTTGCTCTTCTCTCCCGGTGCGGTGGCGCTGTGCAGCGAGTGGCTACAGGATGGGGTGTG  
CGGGCCCGGGGCGCACGGCCCCGACGACGCTTCCGCGCCTCGCTCAGCTGCGTGCTGCC  
CGACTTCTTGAGGGCCGGGCGCCCGCAGCTACGTGGGGGCCTGCTTCGACAGGCTGCT  
CCACCCGGACGCCGTACCCGCCCTTTTCCGCACCGTGCCCGTCTTCACACTGCCCTCCCA  
ACTGCCAGACTTCTGGGGGCCCTGCAGCAGCCTCGCGCCCCGCGTTCCGGGCGGCTCCA  
AGAGAGAGCGGAGCAAGTGTCCCGGGCCCTTCAGCCAGCCCTGGATAGCTACTTCCATCC  
CCCGGGGACTCCCGCGCCGGGACGCGGGGTGGGACCAGGGGCGGGACCTGGGGCGGGGA  
CGGGACTTAAATAAAGGCAGACGCTGTTTTTCTAAAAAAA

379/392

**FIGURE 364**

MPVPWFLLSLALGRSPVVLSSLERLVGPQDATHCSPGLSCRLWDSIDLCLPGDIVPAPGPV  
LAPTHLQTELVLRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNAS  
LQAQVVLVSFOAYPTARCVLLEVQVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSYTQPR  
YEKELNHTQQLPALPWLNVSADGDNVHLVLNVSEEQHFGLSLYWNQVQGPPKPRWHKNLT  
GPQIITLNLHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQNLWQAARLRLTLQSWL  
LDAPCSLPAAEALCWRAPGGDPQCPLVPPLSWENVTVDKVLEFPLLKGHPNLCVQVNSSE  
KLQLQECLWADSLGPLKDDVLLLETRGPQDNRSLEPSGCTSLPSKASTRAARLGEYL  
LQDLQSGQCLQLWDDDLGALWACPMDKYIHKRWALVWLACLLFAAALSLILLKKDHAKG  
WLRLKQDVRSAGAAARGRAALLLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELS  
AQGPVAVFWHAQRRQTLQEGGVVLLFSPGAVALCSEWLQDGVSGPGAHGPHDAFRASLSC  
VLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTVPVFTLPSQLPDLGALQQPRAPRS  
RLQERAEQVSRALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

**Signal sequence:**  
amino acids 1-20

**Transmembrane domain:**  
amino acids 453-475

**N-glycosylation sites:**  
amino acids 118-121, 186-189, 198-201, 211-214, 238-241,  
248-251, 334-337, 357-360, 391-394

**Glycosaminoglycan attachment site:**  
amino acids 583-586

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**  
amino acids 552-555

**N-myristoylation sites:**  
amino acids 107-112, 152-157, 319-324, 438-443, 516-521,  
612-617, 692-697, 696-701, 700-705

380/392

**FIGURE 365**

AATAGAAGTCCTCAGGACGGAGCAGAGGTGGCCGGCGGGCCCGGCTGACTGCGCCTCTGC  
TTTCTTTCCATAACCTTTTCTTTTCGGAATCAGGCTGCTGCGAAGGGTCTAGTTC  
CGGACACTAGGGTGCCCGAACGCGCTGATGCCCCGAGTGCTCGCAGGGCTTCCCGCTAAC  
CATGCTGCCGCCGCCGCCGCCGAGCTGCCTTGGCGCTGCCTGTGCTCCTGCTACTGCT  
GGTGGTGTGACGCCGCCCGACCGGCGCAAGGCCATCCCCAGGCCCAGATTACCTGCG  
GCGCGGCTGGATGCGGCTGCTAGCGGAGGGCGAGGGCTGCGCTCCCTGCCGGCCAGAAGA  
GTGCGCCGCCGCCGCCGGGCTGCCTGGCGGGCAGGGTGCGCGACGCGTGCGGCTGCTGCTG  
GGAATGCGCCAACCTCGAGGGCCAGCTCTGCGACCTGGACCCCAGTGCTCACTTCTACGG  
GCACTGCGGCGAGCAGCTTGAGTGCCGGCTGGACACAGGCGGCGACCTGAGCCGCGGAGA  
GGTGCCGGAACCTCTGTGTGCCTGTGCTTCGACAGTCCGCTCTGCGGGTCCGACGGTCA  
CACCTACTCCAGATCTGCCGCTGCAGGAGGCGGCCCGCGCTCGGCCCGATGCCAACCT  
CACTGTGGCACACCCGGGGCCCTGCGAATCGGGGCCCCAGATCGTGTACATCCATATGA  
CACTTGGAATGTGACAGGGCAGGATGTGATCTTTGGCTGTGAAGTGTTCCTACCCCAT  
GGCCTCCATCGAGTGAGGAAGGATGGCTTGGACATCCAGCTGCCAGGGGATGACCCCCA  
CATCTCTGTGCAGTTTAGGGGTGGACCCAGAGGTTTGAGGTGACTGGCTGGCTGCAGAT  
CCAGGCTGTGCGTCCAGTGATGAGGGCACTTACCGCTGCCTTGGCCGCAATGCCCTGGG  
TCAAGTGGAGGCCCCTGCTAGCTTGACAGTGCTCACACCTGACCAGCTGAACTCTACAGG  
CATCCCCAGCTGCGATCACTAAACCTGGTTCTTGAGGAGGAGGCTGAGAGTGAAGAGAA  
TGACGATTACTACTAGGTCCAGAGCTCTGGCCCATGGGGGTGGGTGAGCGGCTATAGTGT  
TCATCCCTGCTCTTGAAAAGACCTGGAAAGGGGAGCAGGGTCCCTTCATCGACTGCTTTC  
ATGCTGTCAGTAGGGATGATCATGGGAGGCCTATTTGACTCCAAGGTAGCAGTGTGGTAG  
GATAGAGACAAAAGCTGGAGGAGGGTAGGGAGAGAAGCTGAGACCAGGACCGGTGGGGTA  
CAAAGGGGGCCATGCAGGAGATGCCCTGGCCAGTAGGACCTCCAACAGGTTGTTTCCAG  
GCTGGGGTGGGGGCCTGAGCAGACACAGAGGTGCAGGCACCAGGATTCTCCACTTCTTCC  
AGCCCTGCTGGGCCACAGTTCTAACTGCCCTTCTCCAGGCCCTGGTTCTTGCTATTTTC  
CTGGTCCCCAACGTTTATCTAGCTTGTTCGCCCTTCCCCAACTCATCTTCCAGAACTT  
TTCCCTCTCTCCTAAGCCCCAGTTGCACCTACTAACTGCAGTCCCTTTTGCTGTCTGCCG  
TCTTTTGTACAAGAGAGAGAACAGCGGAGCATGACTTAGTTTCAGTGCAGAGAGATTT

381/392

**FIGURE 366**

MLPPPRPAAALALPVLLLLLVLTTPPPTGARPSPGPDYLRRGWMRLLAEGEGCAPCRPEE  
CAAPRGCLAGRVRDACGCCWECANLEGQLCDLDPSAHFYGHCGEQLECRDLDGGDLRGE  
VPEPLCACRSQSPLCGSDGHTYSQICRLQEAARARPDANLTVAHPGPCESGPQIVSHPYD  
TWNVTGQDVIFGCEVFAYPMASIEWRKDGLDIQLPGDDPHISVQFRGGPQRFVETGWLQI  
QAVRPSDEGTYRCLGRNALGQVEAPASLTVLTPDQLNSTGIPQLRSLNLVPEEEAESEEN  
DDYY

**Important features of the protein:****Signal peptide:**

1-30

**Transmembrane domain:**

None

**N-glycosylation site:**

159-163, 183-187, 277-281

**Tyrosine kinase phosphorylation site:**

244-252

**N-myristoylation site:**

52-58, 66-72, 113-119, 249-255

**Kazal-type serine protease inhibitor domain:**

121-168

**Immunoglobulin domain:**

186-255

**Insulin-like growth factor binding proteins:**

53-90

382/392

**FIGURE 367**

AGACGCTACAGGATGGAGCGGGGCGCAGGAGCCAAGCTGCTGCCGCTGCTGCTGCTTCTG  
CGGGCGACTGGTTTACATGTGCACAGACAGATGGCCGGAACGGCTACACGGCGGTATC  
GAAGTGACCAGCGGGGTCCCTGGGGCGACTGGGCCTGGCCTGAGATGTGTCCCGATGGA  
TTCTTCGCCAGCGGGTTCTCGCTCAAGGTGGAGCCTCCCCAAGGCATTCTGGCGACGAC  
ACTGCACTGAATGGGATCAGGCTGCACTGCGCGCGCGGGAACGTCCTAGGCAATACGCAC  
GTGGTAGAGTCCCAGTCTGGAAGCTGGGGCGAATGGAGTGAGCCGCTGTGGTGTGCGCGC  
GGCGCCTACCTAGTGGCTTTCTCGCTTCGCGTGGAGGCACCCACGACCCTCGGTGACAAC  
ACAGCAGCGAACAACGTGCGCTTCGCTGTTGAGACGGCGAGGAACTGCAGGGGCCTGGG  
CTGAGCTGGGGAGACTTTGGAGACTGGAGTGACCATTGCCCAAGGGCGCGTGCGGCCTG  
CAGACCAAGATCCAGGGACCTAGAGGCCTCGGCGATGACACTGCGCTGAACGACGCGCGC  
TTATTCTGCTGCCGCAGTGAACGGCGCCGCCGCCGCTCTCTCCCGGGCCAGGAGGC  
TAGTCCCACCTCTTGCTATTAAAGCTTCTCTGAGTTG

383/392

## FIGURE 368

MERGAGAKLLPLLLLLLRATGFTCAQTDGRNGYTAVIEVTSGGPWGDWAWPEMCPDGFFAS  
GFSLKVEPPQGIPGDDTALNGIRLHCARGNVLGNTHVVESQSGSWGGEWSEPLWCRGGAYL  
VAFSLRVEAPTTLGDNNTAANNVRFRCSDGEELQGPGLSWGDFGDWSDHCPKGACGLQTKI  
QGPRGLGDDTALNDARLFCCRS

Important features of the protein:

Signal peptide:

1-24

Transmembrane domain:

None

N-myristoylation site:

41-47

89-95

156-162

Growth factor and cytokines receptors family signature 2:

103-110

384/392

**FIGURE 369**

GCCAACACTGGCCAAACCTCGGAGACCGTCCTGCGCTCTCTGGAGACGCGCTGTCCGCGC  
CCAGGGTGGTGCCATGTGGGGCGCTCGCCGCTCGTCCGTCTCCTCATCCTGGAACGCCGC  
TTCGCTCCTGCAGCTGCTGCTGGCTGCGCTGCTGGCGGCGGGGGCGAGGGCCAGCGGCGA  
GTACTGCCACGGCTGGCTGGACGCGCAGGGCGTCTGGCGCATCGGCTTCCAGTGTCCTCGA  
GCGCTTCGACGGCGGGCGACGCCACCATCTGCTGCGGCAGCTGCGCGTTGCGCTACTGCTG  
CTCCAGCGCCGAGGCGCGCTGGACCAGGGCGGCTGCGACAATGACCGCCAGCAGGGCGC  
TGGCGAGCCTGGCCGGGCGGACAAAGACGGCCCCGACGGCTCGGCAGTGCCCATCTACGT  
GCCGTTCTCATTTGTTGGCTCCGTGTTTGTGCGCTTTATCATCTTGGGGTCCCTGGTGGC  
AGCCTGTTGCTGCAGATGTCTCCGGCCTAAGCAGGATCCCCAGCAGAGCCGAGCCCCAGG  
GGGTAAACCGCTTGATGGAGACCATCCCCATGATCCCCAGTGCCAGCACCTCCCGGGGGTC  
GTCCTCACGCCAGTCCAGCACAGCTGCCAGTTCAGCTCCAGCGCCAACTCAGGGGCCCCG  
GGCGCCCCAACAAGGTCACAGACCAACTGTTGCTTGCCGGAAGGGACCATGAACAACGT  
GTATGTCAACATGCCCACGAATTTCTCTGTGCTGAACTGTCAGCAGGCCACCCAGATTGT  
GCCACATCAAGGGCAGTATCTGCATCCCCCATACTGCGGTACACGGTGCAGCAGGACTC  
TGTGCCCATGACAGCTGTGCCACCTTTCATGGACGGCCTGCAGCCTGGCTACAGGCAGAT  
TCAGTCCCCCTTCCCTCACACCAACAGTGAACAGAAGATGTACCCAGCGGTGACTGTATA  
ACCGAGAGTCACTGGTGGGTTCTTTACTGAAGGGAGACGAAGGCAGGGGTGGATTTTCG  
AGGTGGAAGT



385/392

**FIGURE 370**

MWGARRSSVSSSWNAASLLQLLLAALLAAGARASGEYCHGWLDAQGVWRIGFQCPERFDG  
GDATICCGSCALRYCCSSAEARLDQGGCDNDRQQGAGEPGRADKDGPDGSAVPIYVPFLI  
VGSVFVAFIILGSLVAACCCRLRPKQDPQQSRAPGGNRLMETIPMIPSASTSRGSSSRQ  
SSTAASSSSSANS GARAPPTRSQTNCCLPEGTMMNVYVNMPTNFSVLNCQQATQIVPHQG  
QYLHPPYVG YTVQHDSVPMTAVPPFMDGLQPGYRQIQSPFPHTNSEQKMPAVTV

**Important features of the protein:**

**Signal peptide:**

1-33

**Transmembrane domain:**

54-78

**N-glycosylation site:**

223-226

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

5-8

**N-myristoylation site:**

3-8, 30-35, 60-65, 86-91, 132-137, 211-216, 268-273

**Prokaryotic membrane lipoprotein lipid attachment site:**

128-138

386/392

## FIGURE 371

CACCAGACAGCACTCCAGCACTCTGTTTGGGGGGCATTCGAAACAGCAAAATCACTCATA  
AAAGGCCAAAAAATTGCAAAAAAATAGTAATAACCAGCATGGCACTAAATAGACCATGA  
AAAGACATGTGTGTGCAGTATGAAAAATTGAGACAGGAAGGCAGAGTGTCTGCTGTCCA  
CCTCAGCTGGGAATGTGCATCAGGCAACTCAAGTTTTTACCACGGCATGTGTCTGTGAA  
TGTCCGCAAAACATTCTCTCTCCCCAGCCTTCATGTGTAACTGGGGATGATGTGGACC  
TGGGCACTGTGGATGCTCCCTTCACTCTGCAAATTCAGCCTGGCAGCTCTGCCAGCTAAG  
CCTGAGAACATTTCCCTGTGTCTACTACTATAGGAAAAATTTAACCTGCACTTGGAGTCCA  
GGAAAGGAAACCAGTTATACCCAGTACACAGTTAAGAGAACTTACGCTTTTGGAGAAAAA  
CATGATAATTGTACAACCAATAGTTCTACAAGTGAAAATCGTGCTTCGTGCTCTTTTTTC  
CTTCCAAGAATAACGATCCCAGATAATTATACCATTTGAGGTGGAAGCTGAAAATGGAGAT  
GGTGTAAATTAAATCTCATATGACATACTGGAGATTAGAGAACATAGCGAAAACCTGAACCA  
CCTAAGATTTTCCGTGTGAAACCAAGTTTTTGGGCATCAAACGAATGATTCAAATTGAATGG  
ATAAAGCTGAGTTGGCGCCTGTTTCATCTGATTAAATAACACACTTCGATTGAGGACA  
GTCAACAGTACCAGCTGGATGGAAGTCAACTTCGCTAAGAACCGTAAGGATAAAAAACCAA  
ACGTACAACCTCACGGGGCTGCAGCCTTTTACAGAATATGTCATAGCTCTGCGATGTGCG  
GTCAAGGAGTCAAAGTTCTGGAGTGACTGGAGCCAAGAAAAAATGGGAATGACTGAGGAA  
GAAGCTCCATGTGGCCTGGAACCTGTGGAGAGTCTGAAACCAGCTGAGGCGGATGGAAGA  
AGGCCAGTGCGGTTGTTATGGAAGAAGGCAAGAGGAGCCCCAGTCCTAGAGAAAACACTT  
GGCTACAACATATGGTACTATCCAGAAAGCAACACTAACCTCACAGAAACAATGAACACT  
ACTAACCAGCAGCTTGAACATGCATCTGGGAGGCGAGAGCTTTTGGGTGTCTATGATTTCT  
TATAATTCTCTTGGGAAGTCTCCAGTGGCCACCCTGAGGATTCCAGCTATTCAAGAAAAA  
TCATTTCAAGTGCATTGAGGTGATGCAGGCCTGCGTTGCTGAGGACCAGCTAGTGGTGAAG  
TGGCAAAGCTCTGCTCTAGACGTGAACACTTGGATGATTGAATGGTTTCCGGATGTGGAC  
TCAGAGCCCCACCACCTTTCTGGGAATCTGTGTCTCAGGCCACGAACCTGGACGATCCAG  
CAAGATAAATTAACCTTTCTGGTGCTATAACATCTCTGTGTATCCAATGTTGCATGAC  
AAAGTTGGCGAGCCATATTCCATCCAGGCTTATGCCAAAGAAGGCGTTCCATCAGAAGGT  
CCTGAGACCAAGGTGGAGAACATTGGCGTGAAGACGGTCACGATCACATGGAAAAGAGATT  
CCCAAGAGTGAGAGAAAGGGTATCATCTGCAACTACACCATCTTTTACCAAGCTGAAGGT  
GGAAAAGGATTCTGTAAGCACGCCCATAGCGAAGTGGAACAAAAACCCCAAGCCCCAGATA  
GATGCTATGGATAGACCTGTTGTAGGCATGGCTCCCCCATCTCATTTGTGACTTGCAACCT  
GGCATGAATCACTTAGCTTCTTTAAATCTCTCTGAAAATGGGGCCAAGAGCACCCACCTT  
TTGGGGTTTTTGGGGGTAAATGAGAGTGAAGTGACAGTACCTGAGAGGAGAGTCCTGAGG  
AAATGGAAGGAGTTGTTATAATTTGTCTGGTTAGGCCCTGAATTGACCTCCCGGGAGCT  
CCCCGACCATCATTCAGGAATGGCGTGCCTGGCTTAAAGAGTGAGGAGGAACAGACCC  
TGTCACCATGACTTCTACTGCCCCTGCCAAATCATGCTTTTGTTTTTTCAGTCCACCTTAT  
CTCCTGACATCTTAAATACTGGGCAAGGCTTGGATTCTTGCTTAGGCTAAATAATTTTTT  
CTTATGGTAAATAACACGTAAAAATTTTTTCCAGTTTAAACATTTGAAAGTGTACAATTT  
AGTGGCATTAGAAGCATTACAAATATTGTGCAACCATCACCATTATTTCCAGAACTCTTC  
TATTTCTGCCCCAAATAGAAGCCCTATACCCATTATTAGTCACTCCCCATTCTCTCCTC  
CCACAGCCCCCTGGCAACTACCAAACTGCTTTGTGTCTCTATGGATTGCCTATTTTGGATA  
TTTCATATACATAGAATCATAAANTAAAAA

387/392

**FIGURE 372**

MCIRQLKFFTTACVCECPQNILSPQPSVCVNLGMMWTWALWMLPSLCKFSLAALPAKPENI  
SCVYYYRKNLCTWSPGKETSYTQYTVKRTYAFGEKHDNCTNSSTSENRASCSFFLPRI  
TIPDNYTIEVEAENG DGVIKSHMTYWRL ENIAKTEPPKIFRVKPV LGIKRMIQIEWIKPE  
LAPVSSDLKYTLRFRTVNSTSWMEVNF AKNRKDKNQTYNLTGLQPFTEYVIALRCAVKES  
KFWSDWSQEKMGMT EEEAPCGLELWRVLKPAEADGRRPVRL LWKKARGAPVLEKTLGYNI  
WYPESNTNLTETMNTTNQQLHLHGGSFWVSMISYNSLGKSPVATLRIPAIQEKSFQC  
IEVMQACVAEDQLVVWKQSSALDVNTWMI EWFPDVDSEPTTLSWESVSQATNWTIQQDKL  
KPFWCYNISVYPMLHDKVGEPYSIQAYAKEGVPSEGPETKVENIGVKTVTITWKEIPKSE  
RKGIIICNYTIFYQAEGGKGFC KHAHSEVEKNPKPQIDAMDRPVVGMAPP SHCDLQPGMNH  
LASLNLSENGAKSTHLLGFWGLNESEVTV PERRVLRKWKELL

**Important features of the protein:**

**Signal peptide:**

1-46

**Transmembrane domain:**

None

**N-glycosylation site:**

59-63, 69-73, 99-103, 103-107, 125-129, 198-202, 215-219, 219-223, 309-313, 315-319, 412-416, 427-431, 487-491, 545-549, 563-567

**N-myristoylation site:**

32-38, 137-143, 483-489, 550-556, 561-567

**Amidation site:**

274-278

**Growth factor and cytokines receptors family signature 1:**

62-75

**Fibronectin type III domain:**

54-144

154-247

388/392

**FIGURE 373**

CCAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCTCGGGGATCCTCTAGAGATCC  
CTCGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTG  
TGGACAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGC  
CCCAGCAAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCAC  
GCCTGGGCTCCAGCAGCATCAGCAGCCCCCAGGACCGGGGGAGGCACAGGTGGCCCCCAC  
CACCCGGAGGAGCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCC  
ACCCAGAGGAGAAGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTG  
CTGCTGATGTGGCTTCTGGTGTGGCAGTGGGCGGCACAGAGCACGCCTACCGGCCCGGC  
CGTTAGGGTGTGTGCTGTCCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTCTGTCAGC  
GTGTGTACCAGCCCTTCTCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAA  
CCATTTATAGGACCGCCTACCGCCGACGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTACG  
CGTGCTGCCCGGCTGGAAGAGGACAGCGGGCTTCTGCGGGCCTGTGGAGCAGCAATAT  
GCCAGCCGCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTG  
CAGGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGGCG  
GCTGTCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTTGGGAGGGGC  
ACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGGCCCCCAGGGTGGCCC  
CCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGG  
TGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGTGGCCCCACTGCACAGCCTGGCCTCGC  
AGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCATCCTTCCAGCAGC  
TCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTCTGGAGGAGCAGCTGGGGTCTCT  
GCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGGCTGGACTGAGCCCCCTCACGC  
CGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCGGGGT  
GACTGAGCGGAAGGCCAGGCAGGGCCTTCTCCTTTCTCCTCCCTTCCCTCGGGAGG  
GTCCCCAGACCCTGGCATGGGATGGGCTGGGATTTTTTTTGTGAATCCACCCCTGGCTAC  
CCCCACCCTGGTTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAGG  
TACGAGTTCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAGGCTGG  
GTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAATGAAACGTGAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTC  
GACCTGCAGAAGCTTGCCCGCATGGCCCAACTTGTATTATGCAGCTTATAATGGTTACAAAT

389/392

**FIGURE 374**

MTDSPPPGHPEEKATPPGGTGHEGLSGGAADVASGVGSGRHRARLPARPLGCVLSRAHGD  
PVSESFVQRVYQPFLTTCDGHRACSTYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGL  
PGACGAATICQPPCRNGGSCVQPGRCRCPAGWRGDTQCSDVDECSARRGGCPQRCINTAGS  
YWCQCWEGHSLSADGTLCPKGGPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEEKQLVL  
APLHSLASQALEHGLPDPGSLLVHSFQQLGRIDSLSEQISFLEEQLGSCSCKKDS

390/392

FIGURE 375

Wholemount In Situ with Prol449 Orthologue

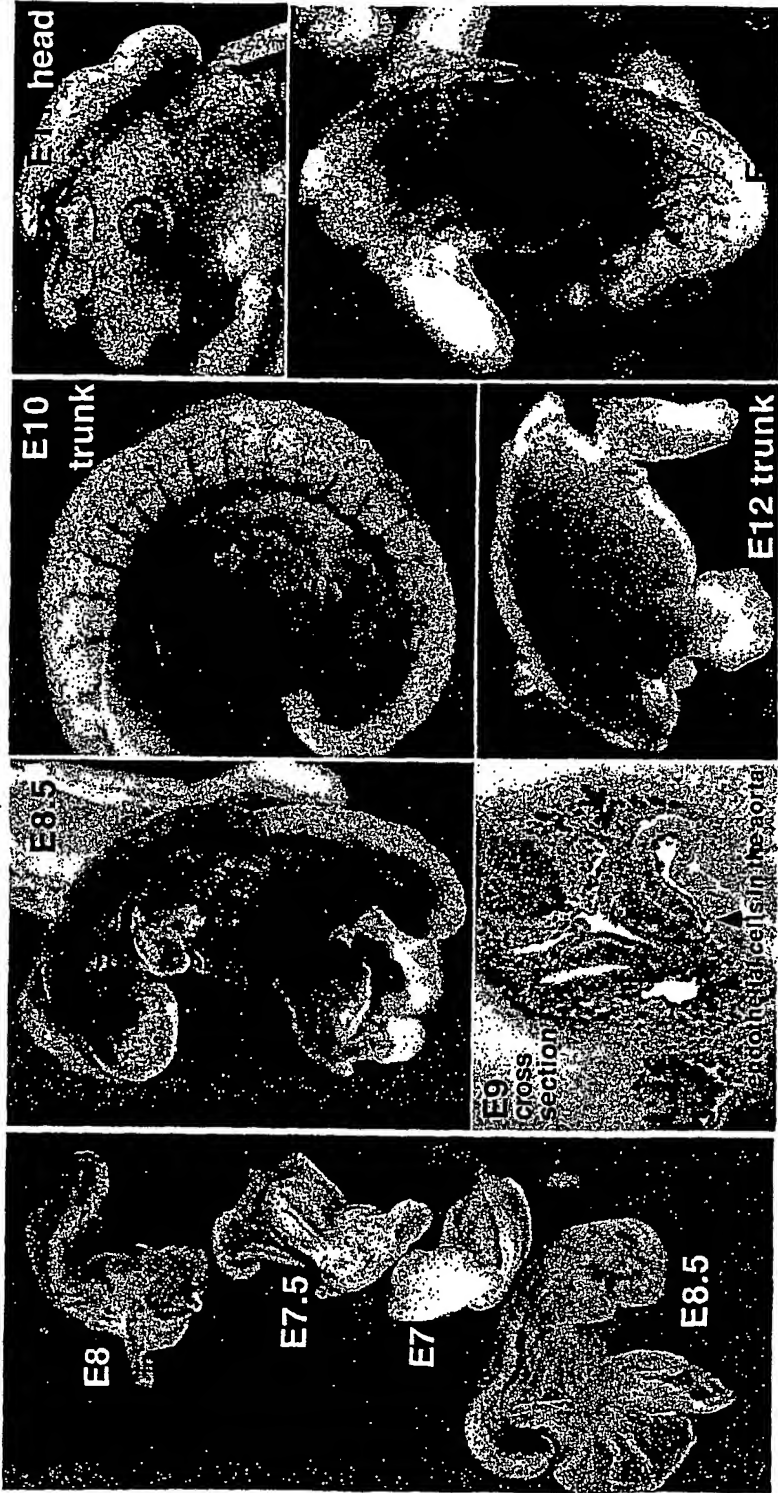
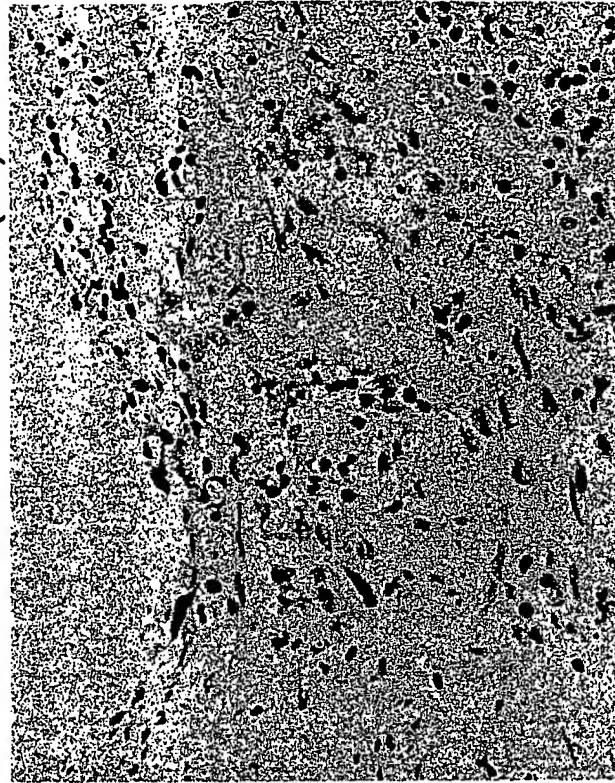
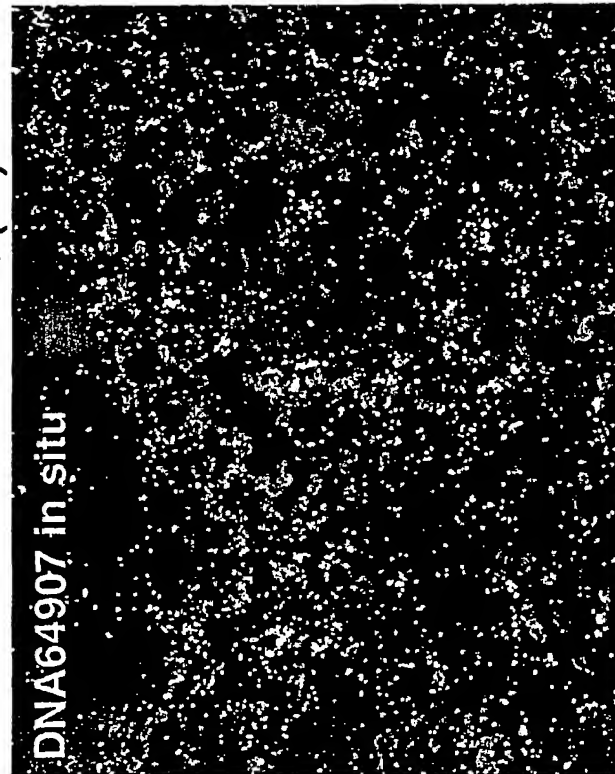


FIGURE 376

PRO1449 is expressed in vasculature of many inflamed and diseased tissues

Human tumor tissue (DF)

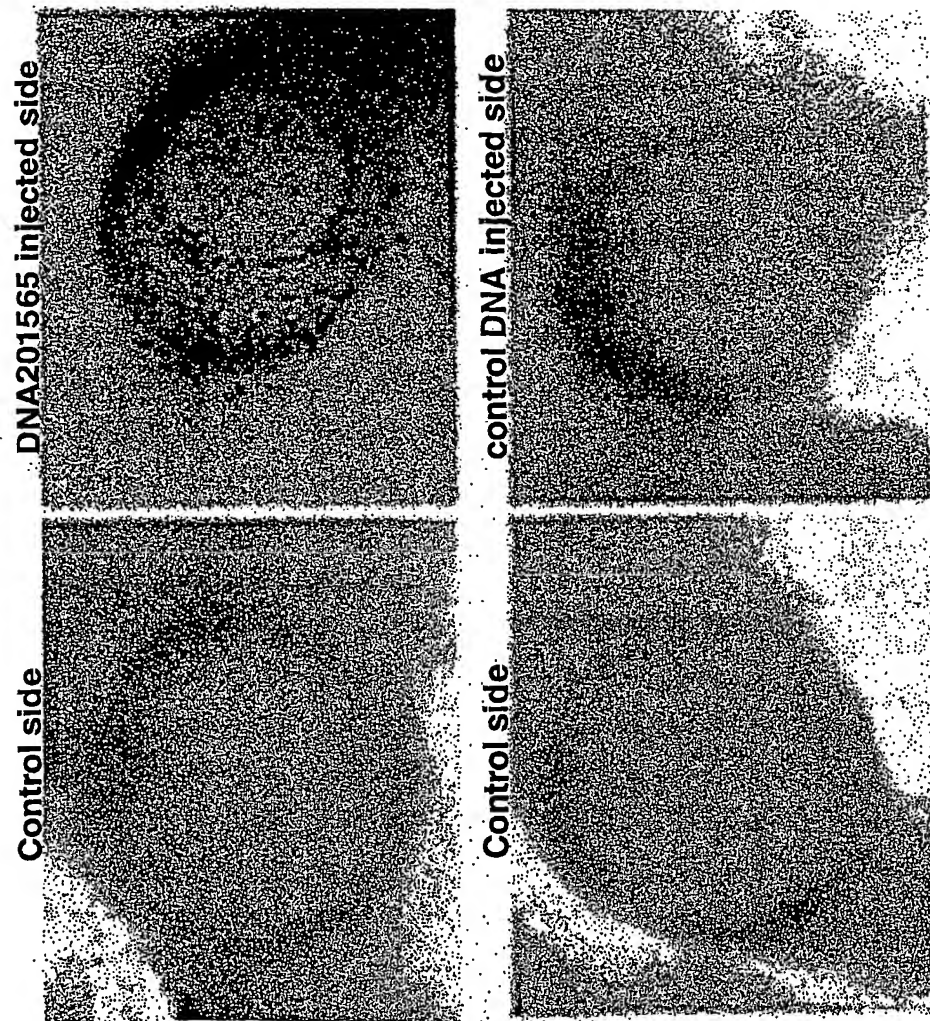
Human tumor tissue (BF)



392/392

## FIGURE 377

Mouse orthologue of PRO1449 induces ectopic vessels in the eyes of chicken embryos





(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
3 January 2002 (03.01.2002)

PCT

(10) International Publication Number  
WO 02/000690 A3

- (51) International Patent Classification<sup>7</sup>: C07K 14/47,  
C12N 15/12, A61K 38/17, C07K 16/18
- (21) International Application Number: PCT/US01/19692
- (22) International Filing Date: 20 June 2001 (20.06.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
- |                |                                |    |
|----------------|--------------------------------|----|
| 60/213,637     | 23 June 2000 (23.06.2000)      | US |
| 60/219,556     | 20 July 2000 (20.07.2000)      | US |
| 60/220,624     | 25 July 2000 (25.07.2000)      | US |
| 60/220,664     | 25 July 2000 (25.07.2000)      | US |
| PCT/US00/20710 | 28 July 2000 (28.07.2000)      | US |
| 60/222,695     | 2 August 2000 (02.08.2000)     | US |
| 09/643,657     | 17 August 2000 (17.08.2000)    | US |
| PCT/US00/23522 | 23 August 2000 (23.08.2000)    | US |
| PCT/US00/23328 | 24 August 2000 (24.08.2000)    | US |
| 60/230,978     | 7 September 2000 (07.09.2000)  | US |
| 60/000,000     | 15 September 2000 (15.09.2000) | US |
| 09/664,610     | 18 September 2000 (18.09.2000) | US |
| 09/665,350     | 18 September 2000 (18.09.2000) | US |
| 60/242,922     | 24 October 2000 (24.10.2000)   | US |
| 09/709,238     | 8 November 2000 (08.11.2000)   | US |
| PCT/US00/30952 |                                |    |
|                | 8 November 2000 (08.11.2000)   | US |
| PCT/US00/30873 |                                |    |
|                | 10 November 2000 (10.11.2000)  | US |
| PCT/US00/32678 |                                |    |
|                | 1 December 2000 (01.12.2000)   | US |
| 09/747,259     | 20 December 2000 (20.12.2000)  | US |
| PCT/US00/34956 |                                |    |
|                | 20 December 2000 (20.12.2000)  | US |
| 09/767,609     | 22 January 2001 (22.01.2001)   | US |
| 09/796,498     | 28 February 2001 (28.02.2001)  | US |
| PCT/US01/06520 |                                |    |
|                | 28 February 2001 (28.02.2001)  | US |
| PCT/US01/06666 | 1 March 2001 (01.03.2001)      | US |
| 09/802,706     | 9 March 2001 (09.03.2001)      | US |
- (71) Applicant (for all designated States except US): GENENTECH, INC. [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): BAKER, Kevin, P. [GB/US]; 14006 Indian Run Drive, Darnestown, MD 20878 (US). FERRARA, Napoleone [US/US]; #704, 2090 Pacific Avenue, San Francisco, CA 94109 (US). GERBER, Hanspeter [CH/US]; #5, 1121 Tennessee Street, San Francisco, CA 94107 (US). GERRITSEN, Mary, E. [CA/US]; 541 Parrott Drive, San Mateo, CA 94402 (US). GODDARD, Audrey [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). GODOWSKI, Paul, J. [US/US]; 25 Orange Court, Hillsborough, CA 94010 (US). GURNEY, Austin, L. [US/US]; 1 Debbie Lane, Belmont, CA 94002 (US). HILLAN, Kenneth, J. [GB/US]; 64 Seward Street, San Francisco, CA 94114 (US). MARSTERS, Scot, A. [US/US]; 990 Cherry Street, San Carlos, CA 94070 (US). PAN, James [CA/US]; 2705 Coronet Boulevard, Belmont, CA 94002 (US). PAONI, Nicholas, F. [US/US]; 1756 Terrace Drive, Belmont, CA 94002 (US). STEPHAN, Jean-Philippe, F. [FR/US]; 320 C Lansdale Avenue, Millbrae, CA 94030 (US). WATANABE, Colin, K. [US/US]; 128 Corliss Drive, Moraga, CA 94556 (US). WILLIAMS, P., Mickey [US/US]; 509 Alto Avenue, Half Moon Bay, CA 94019 (US). WOOD, William, I. [US/US]; 35 Southdown Court, Hillsborough, CA 94010 (US). YE, Weilan [CN/US]; 119 Barkentine Street, Foster City, CA 94404 (US).

[Continued on next page]

(54) Title: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS

(57) Abstract: Compositions and methods are disclosed for stimulating or inhibiting angiogenesis and/or cardiovascularization in mammals, including humans. Pharmaceutical compositions are based on polypeptides or antagonists thereto that have been identified for one or more of these uses. Disorders that can be diagnosed, prevented, or treated by the compositions herein include trauma such as wounds, various cancers, and disorders of the vessels including atherosclerosis and cardiac hypertrophy. In addition, the present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

WO 02/000690 A3



(74) Agents: AGARWAL, Atulya, R. et al.; c/o GENENTECH, INC., MS49, 1 DNA Way, South San Francisco, CA 94080-4990 (US).

patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

**Published:**

— with international search report

(88) Date of publication of the international search report:  
13 March 2003

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/19692

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C07K14/47 A61K38/17 C07K16/18

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EP0-Internal

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,L	WO 99 46281 A (BAKER KEVIN P ;CHEN JIAN (US); GENENTECH INC (US); GURNEY AUSTIN ( ) 16 September 1999 (1999-09-16) see PR0181	1-19
X,L	US 5 968 744 A (CORLEY-NEIL C ET AL) 19 October 1999 (1999-10-19) L: priority. the whole document	1-19
X	WO 99 50405 A (GENETICS INST) 7 October 1999 (1999-10-07) see pk65_4 and seq.ID's 11 and 12	1-19
X	WO 99 43802 A (KATO SEISHI ;KIMURA TOMOKO (JP); PROTEGENE INC (JP); SEKINE SHINGO) 2 September 1999-(1999-09-02)- see clone HP02239.	1-19
-/--		

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

## \* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&amp;" document member of the same patent family

Date of the actual completion of the international search

21 August 2002

Date of mailing of the international search report

29. 11. 2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3018

Authorized officer

Smalt, R

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/19692

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 10363 A (WEI YING FEI ;HUMAN GENOME SCIENCES INC (US); ROSEN CRAIG A (US);) 4 March 1999 (1999-03-04) see secreted protein 24 ---	1-19
X	WO 99 33979 A (CHIRON CORP) 8 July 1999 (1999-07-08) the whole document ---	1-19
A	WO 95 29242 A (CHILDRENS MEDICAL CENTER) 2 November 1995 (1995-11-02) the whole document ---	
A	EP 0 335 243 A (HARVARD COLLEGE) 4 October 1989 (1989-10-04) the whole document ---	
A	FOLKMAN J: "CLINICAL APPLICATIONS OF RESEARCH ON ANGIOGENESIS" NEW ENGLAND JOURNAL OF MEDICINE, THE, MASSACHUSETTS MEDICAL SOCIETY, WALTHAM, MA, US, vol. 333, no. 26, 28 December 1995 (1995-12-28); pages 1757-1763, XP000654557 ISSN: 0028-4793 the whole document ---	
A	KLGSBRUN M: "REGULATORS OF ANGIOGENESIS: STIMULATORS, INHIBITORS, AND EXTRACELLULAR MATRIX" JOURNAL OF CELLULAR BIOCHEMISTRY, LISS, NEW YORK, NY, US, vol. 47, no. 3, 1 November 1991 (1991-11-01), pages 199-200, XP000616588 ISSN: 0730-2312 the whole document ---	
P,X, L	WO 00 53756 A (BOTSTEIN DAVID;BAKER KEVIN P ; GENENTECH INC (US); ASHKENAZI AVI J) 14 September 2000 (2000-09-14) see seq.ID.322; L: priority. ---	1-19
P,X, L	WO 00 53751 A (BAKER KEVIN P ;GENENTECH INC (US); NAPIER MARY A (US); ASHKENAZI A) 14 September 2000 (2000-09-14) the whole document ---	1-21,23, 27-29, 31,32, 34,35
P,X	WO 01 00828 A (CORIXA CORP ;FANGER GARY R (US); RETTER MARC W (US); WANG TONGTONG) 4 January 2001 (2001-01-04) see seq.ID.327 ---	1-21,23, 27-29, 31,32, 34,35
	---	

-/--

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/19692

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E,L	WO 02 08284 A (FERRARA NAPOLEONE ;STEPHAN JEAN PHILIPPE F (US); WILLIAMS P MICKEY) 31 January 2002 (2002-01-31) See seq.ID's 1 and 120. L: priority. -----	1-35,37, 41

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US 01/19692

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:  
Although claims 20-34, 37 and 41 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☒ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:  
see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
1-35, 37, 41 all partially

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

## Continuation of Box I.2

Present claims 20-28,31,34,35,37, and 41 relate to products defined by reference to a desirable characteristic or property, namely having (ant)agonistic activity towards the protein(s) of claim 1.

The application provides no support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for any such products. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the product by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search impossible. Consequently, the search of said claims, in as far as the (ant)agonists are concerned, has been carried out for those aspects which appear to be clear, supported and disclosed, namely those parts relating to antibodies directed against the protein(s) of claim 1.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1: 1-35, 37, 41 all partially

Nucleic acid with at least 80% identity to seq.ID.1 and protein encoded thereby, vector, host cell, method for producing the protein, chimeric protein, antibody, and pharmaceutical compositions.

Inventions 2-187: 1-42, all partially,  
and as far as applicable

Subject matter as defined for invention 1. above, but limited to the respective nucleic acid sequences 3-373 (odd numbers) and the polypeptides encoded thereby (seq.ID's 4-374, even numbers).

For the sake of conciseness, the first subject matter is explicitly defined, the other subject matters are defined by analogy thereto.



## INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 01/19692

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9946281	A	16-09-1999	US 6391311 B1	21-05-2002
			AU 3072199 A	27-09-1999
			CA 2321677 A1	16-09-1999
			US 2002127584 A1	12-09-2002
			US 2002094416 A1	18-07-2002
			WO 9946281 A2	16-09-1999
			US 2002156006 A1	24-10-2002
			US 2002169284 A1	14-11-2002
			ZA 9901937 A	17-04-2000
			EP 1241187 A2	18-09-2002
			AU 3075099 A	11-10-1999
			CA 2322792 A1	23-09-1999
			EP 1064382 A2	03-01-2001
			WO 9947677 A2	23-09-1999
			US 6455283 B1	24-09-2002
			EP 1241183 A2	18-09-2002
			US 2002072496 A1	13-06-2002
			US 2002142961 A1	03-10-2002
			US 2002072067 A1	13-06-2002
			US 2002072092 A1	13-06-2002
			US 2002072497 A1	13-06-2002
			US 2002103125 A1	01-08-2002
			US 2002123463 A1	05-09-2002
			US 2002132252 A1	19-09-2002
			US 2002127576 A1	12-09-2002
			US 2002132253 A1	19-09-2002
			US 2002160384 A1	31-10-2002
			AU 735081 B2	28-06-2001
			AU 1532499 A	15-06-1999
			CA 2309358 A1	03-06-1999
			EP 1241180 A2	18-09-2002
			EP 1032667 A2	06-09-2000
			WO 9927098 A2	03-06-1999
			US 6410708 B1	25-06-2002
			EP 1241185 A2	18-09-2002
			EP 1249491 A2	16-10-2002
			EP 1251139 A2	23-10-2002
			US 2002161199 A1	31-10-2002
			EP 1241179 A2	18-09-2002
			EP 1241186 A2	18-09-2002
			AU 741133 B2	22-11-2001
			AU 3757099 A	08-11-1999
			CA 2324297 A1	28-10-1999
			EP 1071773 A1	31-01-2001
			JP 2002512032 T	23-04-2002
			WO 9954467 A1	28-10-1999
			US 2002147322 A1	10-10-2002
			EP 1241251 A2	18-09-2002
			US 2002119118 A1	29-08-2002
-----				
US 5968744	A	19-10-1999	US 2002103342 A1	01-08-2002
			US 6348576 B1	19-02-2002
-----				
WO 9950405	A	07-10-1999	AU 3457299 A	18-10-1999
			CA 2324492 A1	07-10-1999
			EP 1068307 A1	17-01-2001
			JP 2002509722 T	02-04-2002
			WO 9950405 A1	07-10-1999

## INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 01/19692

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 9943802	A	02-09-1999	AU	2549399 A	15-09-1999
			CA	2322711 A1	02-09-1999
			EP	1056847 A2	06-12-2000
			WO	9943802 A2	02-09-1999
			JP	2002506615 T	05-03-2002
WO 9910363	A	04-03-1999	AU	8921598 A	16-03-1999
			CA	2302387 A1	04-03-1999
			EP	1017707 A1	12-07-2000
			JP	2002532054 T	02-10-2002
			WO	9910363 A1	04-03-1999
			US	6410709 B1	25-06-2002
			US	2001016647 A1	23-08-2001
WO 9933979	A	08-07-1999	AU	1929599 A	19-07-1999
			EP	1042470 A2	11-10-2000
			JP	2002511231 T	16-04-2002
			WO	9933979 A2	08-07-1999
			US	2002034800 A1	21-03-2002
WO 9529242	A	02-11-1995	US	5639725 A	17-06-1997
			US	5792845 A	11-08-1998
			AU	692865 B2	18-06-1998
			AU	2461795 A	16-11-1995
			BR	9507479 A	16-09-1997
			CA	2188813 A1	02-11-1995
			CN	1149319 A	07-05-1997
			CZ	9603122 A3	14-05-1997
			EP	0758390 A1	19-02-1997
			HU	76095 A2	30-06-1997
			JP	9512173 T	09-12-1997
			NZ	285501 A	26-02-1998
			WO	9529242 A1	02-11-1995
			US	5885795 A	23-03-1999
			US	5733876 A	31-03-1998
			US	5776704 A	07-07-1998
			ZA	9503419 A	11-01-1996
EP 0335243	A	04-10-1989	US	4900673 A	13-02-1990
			AT	96844 T	15-11-1993
			AU	621358 B2	12-03-1992
			AU	3433189 A	16-10-1989
			CA	1331356 A1	09-08-1994
			DE	68910354 D1	09-12-1993
			DE	68910354 T2	03-03-1994
			DK	233390 A	26-11-1990
			EP	0335243 A2	04-10-1989
			IE	890954 L	28-09-1989
			JP	3503641 T	15-08-1991
			PT	90129 A	10-11-1989
			WO	8909277 A1	05-10-1989
			ZA	8902270 A	24-12-1991
WO 0053756	A	14-09-2000	AU	1748200 A	19-06-2000
			AU	1749900 A	12-07-2000
			AU	2192800 A	12-07-2000
			AU	2224800 A	28-09-2000
			AU	2399300 A	28-09-2000

## INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 01/19692

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 0053756	A		AU 2495200 A	28-09-2000
			AU 2596700 A	28-09-2000
			AU 2600800 A	28-09-2000
			AU 2879400 A	31-07-2001
			AU 2883600 A	28-09-2000
			AU 3072199 A	27-09-1999
			AU 3107700 A	28-09-2000
			AU 3381600 A	28-09-2000
			AU 3514400 A	28-09-2000
			CA 2321677 A1	16-09-1999
			EP 1173563 A1	23-01-2002
			EP 1159419 A1	05-12-2001
			EP 1220905 A2	10-07-2002
			EP 1135485 A2	26-09-2001
			EP 1141284 A2	10-10-2001
			EP 1159422 A1	05-12-2001
			EP 1141289 A2	10-10-2001
			US 2002119130 A1	29-08-2002
			WO 0053753 A2	14-09-2000
			WO 0053754 A1	14-09-2000
			WO 0053755 A2	14-09-2000
			WO 0153486 A1	26-07-2001
			WO 0053756 A2	14-09-2000
			WO 0053757 A2	14-09-2000
			WO 0053758 A2	14-09-2000
			US 2002127584 A1	12-09-2002
			US 2002094416 A1	18-07-2002
			WO 0032221 A2	08-06-2000
			WO 0053750 A1	14-09-2000
			WO 0037638 A2	29-06-2000
			WO 0037640 A2	29-06-2000
			WO 0053751 A1	14-09-2000
			WO 0053752 A2	14-09-2000
			US 2002119118 A1	29-08-2002
			US 2002161199 A1	31-10-2002
			US 2002156006 A1	24-10-2002
			US 2002169284 A1	14-11-2002
			US 2002072496 A1	13-06-2002
			US 2002142961 A1	03-10-2002
			US 2002072067 A1	13-06-2002
			US 2002072092 A1	13-06-2002
			US 2002072497 A1	13-06-2002
			US 2002103125 A1	01-08-2002
-----				
WO 0053751	A	14-09-2000	AU 1748200 A	19-06-2000
			AU 1749900 A	12-07-2000
			AU 2192800 A	12-07-2000
			AU 2224800 A	28-09-2000
			AU 2399300 A	28-09-2000
			AU 2495200 A	28-09-2000
			AU 2596700 A	28-09-2000
			AU 2600800 A	28-09-2000
			AU 2879400 A	31-07-2001
			AU 2883600 A	28-09-2000
			AU 3072199 A	27-09-1999
			AU 3107700 A	28-09-2000
			AU 3381600 A	28-09-2000
			AU 3514400 A	28-09-2000

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International Application No  
PCT/US 01/19692

Patent document cited in search report	Publication date	Patent family member(s)	Publication date	
WO 0053751	A	CA 2321677 A1	16-09-1999	
		EP 1173563 A1	23-01-2002	
		EP 1159419 A1	05-12-2001	
		EP 1220905 A2	10-07-2002	
		EP 1135485 A2	26-09-2001	
		EP 1141284 A2	10-10-2001	
		EP 1159422 A1	05-12-2001	
		EP 1141289 A2	10-10-2001	
		US 2002119130 A1	29-08-2002	
		WO 0053753 A2	14-09-2000	
		WO 0053754 A1	14-09-2000	
		WO 0053755 A2	14-09-2000	
		WO 0153486 A1	26-07-2001	
		WO 0053756 A2	14-09-2000	
		WO 0053757 A2	14-09-2000	
		WO 0053758 A2	14-09-2000	
		US 2002127584 A1	12-09-2002	
		US 2002094416 A1	18-07-2002	
		WO 0032221 A2	08-06-2000	
		WO 0053750 A1	14-09-2000	
		WO 0037638 A2	29-06-2000	
		WO 0037640 A2	29-06-2000	
		WO 0053751 A1	14-09-2000	
		WO 0053752 A2	14-09-2000	
		US 2002119118 A1	29-08-2002	
		US 2002161199 A1	31-10-2002	
		US 2002156006 A1	24-10-2002	
		US 2002169284 A1	14-11-2002	
		US 2002072496 A1	13-06-2002	
		US 2002142961 A1	03-10-2002	
		US 2002072067 A1	13-06-2002	
		US 2002072092 A1	13-06-2002	
		US 2002072497 A1	13-06-2002	
		US 2002103125 A1	01-08-2002	
WO 0100828	A	04-01-2001	AU 5902700 A	31-01-2001
			BR 0012069 A	27-08-2002
			CN 1367829 T	04-09-2002
			EP 1196565 A2	17-04-2002
			HU 3201757 A2	28-09-2002
			NO 20016427 A	25-02-2002
			WO 0100828 A2	04-01-2001
			US 2002168637 A1	14-11-2002
			TR 200200691 T2	21-10-2002
WO 0208284	A	31-01-2002	US 2002042367 A1	11-04-2002
			US 2002010137 A1	24-01-2002
			US 2002106743 A1	08-08-2002
			US 2002052027 A1	02-05-2002
			US 2002058309 A1	16-05-2002
			US 2002102723 A1	01-08-2002
			AU 6531101 A	17-12-2001
			AU 7197301 A	05-02-2002
			AU 7315001 A	30-01-2002
			AU 7885201 A	08-01-2002
			WO 0193983 A1	13-12-2001
			WO 0200690 A2	03-01-2002
			WO 0208288 A2	31-01-2002

## INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 01/19692

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0208284	A	WO 0208284 A2	31-01-2002
		AU 1916701 A	16-07-2001
		AU 2055401 A	12-06-2001
		AU 6391000 A	19-02-2001
		AU 6802801 A	24-09-2001
		AU 6901800 A	12-06-2001
		EP 1244784 A2	02-10-2002
		EP 1250426 A2	23-10-2002
		WO 0140464 A1	07-06-2001
		WO 0149715 A2	12-07-2001
		WO 0140466 A2	07-06-2001
		WO 0168848 A2	20-09-2001
		US 2002127584 A1	12-09-2002
		US 2002094416 A1	18-07-2002
		US 2002160374 A1	31-10-2002
		US 2002146709 A1	10-10-2002
		US 2002132240 A1	19-09-2002
		US 2002142958 A1	03-10-2002
		US 2002150976 A1	17-10-2002
		US 2002132981 A1	19-09-2002
		US 2002165143 A1	07-11-2002
		US 2002156004 A1	24-10-2002
		US 2002142419 A1	03-10-2002
		US 2002102647 A1	01-08-2002
		US 2002110859 A1	15-08-2002
		US 2002142959 A1	03-10-2002
		US 2002115145 A1	22-08-2002
		US 2002132768 A1	19-09-2002
		US 2002127643 A1	12-09-2002
		US 2002156006 A1	24-10-2002
		US 2002169284 A1	14-11-2002
		US 2002072496 A1	13-06-2002
		US 2002142961 A1	03-10-2002
		US 2002072067 A1	13-06-2002
		US 2002072092 A1	13-06-2002
		US 2002072497 A1	13-06-2002
		US 2002103125 A1	01-08-2002